

Calibration de modèles et couverture de produits dérivés

Peter Tankov

▶ **To cite this version:**

Peter Tankov. Calibration de modèles et couverture de produits dérivés. DEA. Calibration de modèles et couverture de produits dérivés, Master "Modélisation Aléatoire", Université Paris-Diderot (Paris 7), 2006, pp.143. cel-00664993

HAL Id: cel-00664993
https://cel.hal.science/cel-00664993

Submitted on 31 Jan 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Calibration de Modèles et Couverture de Produits Dérivés

Peter TANKOV
Université Paris-Diderot (Paris VII)
tankov@math.univ-paris-diderot.fr

Edition 2009, dernière m.à.j. le 16 janvier 2011

Ce document est mis à disposition sous un contrat Creative Commons
(pas d'Utilisation Commerciale — pas de Modification)

La dernière version est disponible à l'adresse

`www.math.jussieu.fr/~tankov/MA/`

Table de matières

1 Les marchés de produits dérivés **7**
1.1 Historique . . . 7
1.2 Fonctionnement des marchés d'options . . . 7
1.3 Différents types de sous-jacents . . . 7
1.4 Les options européennes . . . 8
1.5 Les options américaines . . . 10
1.6 Les options exotiques . . . 10

I Modèles **11**

2 Le modèle de Black et Scholes et la couverture en delta **13**
2.1 Le modèle de Black et Scholes . . . 13
2.1.1 Portefeuille autofinançant . . . 13
2.1.2 Evaluation risque-neutre . . . 14
2.1.3 Formule de Black et Scholes . . . 14
2.1.4 Option sur un actif versant des dividendes . . . 15
2.1.5 Exemples de couverture en delta . . . 15
2.1.6 Analyse de l'erreur de couverture . . . 15
2.1.7 Couverture en gamma . . . 19
2.2 L'utilisation du modèle de Black-Scholes dans les marchés d'options . . . 20
2.2.1 La couverture en delta comme une stratégie de trading . . . 20
2.2.2 Robustesse de la formule de Black-Scholes . . . 21
2.3 La volatilité implicite . . . 22
2.3.1 Rôle de la volatilité implicite dans les marchés d'options . . . 23
2.3.2 Prise en compte des dividendes dans le calcul de la volatilité implicite . . . 24
2.4 Complément: formule de Garman-Kohlhagen et cotation d'options sur taux de change . . . 24

3 Les modèles à volatilité locale et la diffusion implicite **27**
3.1 Modèles à volatilité locale . . . 27
3.2 Arbre trinomial de pricing . . . 29
3.3 Arbres implicites . . . 32
3.4 Diffusion implicite de Dupire . . . 34

4 Les modèles à volatilité stochastique **41**
4.1 Equations de pricing et de couverture . . . 41
4.2 Comportement asymptotique de volatilité implicite . . . 45
4.3 Estimation de volatilité . . . 49
4.4 Swaps de variance . . . 53
4.5 Paramétrisation de Heston . . . 55

5 Les modèles avec sauts **61**
5.1 Introduction aux processus de Lévy . . . 61
5.2 Structure des trajectoires d'un processus de Lévy . . . 66
5.3 Modèles exponentielle-Lévy . . . 69
5.4 Bases du calcul stochastique pour les processus avec sauts . . . 71
5.5 Exponentielle stochastique d'un processus avec sauts . . . 76
5.6 Couverture dans les modèles avec sauts . . . 80
5.7 Valorisation d'options dans les modèles exp-Lévy . . . 85
5.8 Méthode de transformée de Fourier pour la valorisation d'options européennes . . . 87
5.9 Calibration de modèles exp-Lévy . . . 89
5.10 Limites et extensions des processus de Lévy . . . 93

II Outils **95**

6 Régularisation des problèmes mal posés **97**
6.1 Introduction . . . 97
6.2 Problèmes mal posés linéaires . . . 99
6.3 Régularisation de problèmes mal posés . . . 101
6.4 Régularisation de Tikhonov pour les problèmes linéaires . . . 102
6.5 Régularisation du problème de reconstruction de la courbe de taux d'intérêt . . . 105
6.5.1 Méthodes classiques . . . 106
6.5.2 Méthode de régularisation . . . 107

7 Régularisation de problèmes de calibration I: problèmes linéaires **111**
7.1 Reconstruction de la densité risque-neutre . . . 111
7.1.1 Application de la régularisation de Tikhonov à la reconstruction de la SPD . . . 112

8 Régularisation de problèmes de calibration II: problèmes non-linéaires **117**
8.1 Régularisation de Tikhonov pour les problèmes non-linéaires . . . 117
8.2 Régularisation du problème de calcul de la volatilité locale . . . 118

9 Méthodes numériques pour la calibration de modèles **123**
9.1 Algorithmes d'optimisation . . . 123
9.1.1 Optimisation: rappels théoriques . . . 123
9.1.2 Problèmes d'optimisation en finance . . . 124
9.1.3 Minimisation de fonctions convexes . . . 124

A Quelques rappels sur le calcul stochastique et les processus de Markov **129**

B Examens des années précedentes **131**
B.1 Master Modélisation Aléatoire, examen du 27 avril 2006 . . . 131
B.2 Master Modélisation Aléatoire, examen du 10 mai 2007 . . . 133
B.3 Master Modélisation Aléatoire, examen du 14 mai 2008 . . . 136
B.4 Master Modélisation Aléatoire, examen du 7 mai 2009 . . . 138
B.5 Master Probabilités et Finance, examen du 27 mars 2009 . . . 139

Bibliographie **141**

Introduction

> If you want to know the value of a security, use the price of another security that is as similar to it as possible. All the rest is modeling.
> *Emmanuel Derman*

En paraphrasant la celebre formule d'Emmanuel Derman, donnée en épigraphe, on peut dire que pour valoriser un produit complexe, il faut tout d'abord trouver des actifs liquides qui lui ressemblent, et dont les prix sont cotés sur le marché et fiables. Ensuite, l'approche le plus souvent utilisée consiste à choisir un *modèle*, dont les paramètres seront estimés sur les prix des produits liquides (calibration proprement dite). Ce modèle calibré peut maintenant être utilisé pour valoriser le produit liquide. Différents modèles vont, a priori, donner des prix différents, et on peut legitimement se demander, quelles sont les bornes imposées sur le prix du produit exotique uniquement par l'absence d'arbitrage. Bien que pour certains produits, on peut obtenir ces bornes en ne faisant pas ou très peu d'hypothèses sur le modèle, voir par exemple la section sur les swaps de variance, pour la plupart de pay-offs cela s'avère impossible. En pratique, on utilise donc presque toujours un modèle pour faire le lien entre les prix des options liquides et celui du produit exotique. Le cours de caliration est donc avant tout un cours de différents modèles utilisés dans le marché de produits dérivés. Dans ces notes, on decouvrira avant tout les modèles de volatilité (volatilité locale, volatilité stochastique), mais aussi les modèles prenant en compte le risque de saut (processus de Lévy).

La calibration utilise les prix d'options cotées sur le marché. On va donc comprendre le fonctionnement des marchés et les méthodes de cotation. Les prix cotés sont représentés sous la forme d'une *smile* de volatilité; un cours de calibration est donc un cours de modélisation de smile.

Du point de vue théorique, le problème de calibration, c'est-à-dire, reconstruction des paramètres du modèle à partir des observations de prix est un problème inverse mal posé. Inverse car dans tout modèle on sait facilement résoudre le problème de pricing d'options européennes (sinon ce modèle ne sera jamais utilisé), et pour la calibration on cherche à inverser ce processus. Mal posé car cette inversion est très souvent instable, et conduit à l'amplification des erreurs de données dans la solution. On aura donc besoin de la théorie de problèmes inverses mal posés. Mathématiquement, un problème inverse mal posé consiste à résoudre l'équation

$$Tx = y, \tag{1}$$

où T est un opérateur (linéaire ou non) dont l'inverse T^{-1} n'est pas continu. Typiquement, y est le vecteur de données observées sur le marché et x est le vecteur des paramètres du modèle qu'on cherche à calibrer. Dans (1), T^{-1} peut être connu explicitement, où peut être calculé itérativement en minimisant $\|Tx - y\|^2$, ceci ne change rien à la nature mal posé du problème. Si le problème est mal posé une petite erreur sur les données (y) peut conduire à une très grande erreur sur la solution (x), rendant impossible toute résolution en utilisant les données observées (car les observations contient toujours un certain bruit). La théorie de régularisation de problèmes inverses mal posés [29] ne permet pas de transformer un problème mal posé en un problème bien posé, il y aura toujours une perte de précision, mais elle permet de récupérer le maximum d'information sur x à partir de y de façon stable.

En statistique on cherche souvent à résoudre un problème similaire, par exemple, celui d'estimer x dans

$$Tx = y$$

à partir d'une observation bruitée $\tilde{y} = y + \varepsilon$. La différence est qu'en statistique on considère les observations comme des variables aléatoires, qui contient toujours un niveau de bruit assez important, et on cherche à construire des estimateurs convergents lorsque le nombre d'observations tend vers l'infini. Cela n'a pas de sens de "calibrer" les observations, le critère étant la qualité d'estimation du paramètre inconnu.

En calibration, on regarde les prix observés sur le marché comme les 'vrais prix'. Ils peuvent contenir un peu de bruit, à cause des fourchettes bid-ask, mais on cherche à les calibrer avec une grande précision, sinon on nous arbitre. Nous ne pouvons pas faire tendre le nombre d'observations vers l'infini, car il est fixé par le marché, la limite qui a un sens c'est la limite "niveau de bruit d'observation tend vers zéro", ce qui correspond à un marché très liquide.

Les méthodes numériques dont on a souvent besoin en calibration et qu'on étudiera dans ce cours sont

- Méthodes de pricing permettant de calculer les prix de plusieurs options en même temps (méthodes d'arbres, EDP forward, méthodes de transformée de Fourier).
- Méthodes de régularisation (pour la résolution stable de problèmes inverses mal posés).
- Méthodes d'optimisation (algorithmes classiques, optimisation par Monte Carlo).

Finalement, *last but not least*, l'objectif de ce cours est d'apprendre comment choisir un modèle adapté à un produit donné et une méthode de calibration adapté à ce modèle. Le choix du modèle ne peut pas s'effectuer dans l'absolu, car il n'existe pas de modèle qui prendrait en compte tous les risques de tous les produits qu'on peut imaginer. Pour différents produits, différents risques sont importants, et on va choisir le modèle qui prend en compte ces risques et en neglige d'autres, qui sont jugés peu importants pour le produit en question.

Le choix d'un modèle et d'une méthode de calibration en vue de calculer le prix d'un produit exotique s'effectue alors autour des trois questions fondamentales:

Risque de modèle. Il s'agit de déterminer, dans quelle mesure le prix final du produit exotique est expliqué par le modèle utilisé, et dans quelle mesure il est expliqué par les prix des instruments de calibration utilisés. Idéalement, le prix d'un exotique doit être complètement déterminé par les instruments de calibration et ne doit pas avoir une dépendance importante par rapport au modèle utilisé. Si tel n'est pas le cas, on parle du *risque de modèle*. Par exemple, si on utilise un modèle calibré aux options européennes d'échéance T pour valoriser un autre pay-off européen ayant la même échéance T, le risque du modèle est petit. Si, au contraire, on utilise le même modèle et les mêmes produits de calibration pour valoriser un produit "path-dependent", le risque du modèle sera plus grand.

Le risque du modèle est en grande partie déterminé par les instruments de calibration utilisés: pour le diminuer il faut choisir des instruments de calibration naturels pour le produit exotique en question, (exemple: swaptions co-terminales pour valoriser une swaption bermudéenne) et il faut que l'ensemble d'instruments de calibration soit suffisamment complet.

Adéquation du modèle choisi aux instruments de calibration. Ici, on cherche à répondre à la question suivante: est-ce que le modèle est suffisamment riche pour bien reproduire les prix de tous les instruments de calibration? Parfois, on a cette propriété par construction, dans d'autres situations on n'a pas cette propriété a priori, et il faut vérifier a posteriori qu'elle est satisfaite avec une précision suffisante.

Validité et stabilité de l'algorithme de calibration. En plus d'être mathématiquement correct, l'algorithme de calibration utilisé doit être numériquement stable, c'est-à-dire, les petits changements des prix des instruments de calibration ne doivent pas conduire à des grands changements au niveau des paramètres calibrés. Par exemple, si l'algorithme fait intervenir une minimisation numérique, la fonctionnelle à minimiser doit être convexe ou en tout cas ne doit pas avoir de minimums locaux, qui conduisent souvent à des instabilités numériques.

Chapitre 1

Les marchés de produits dérivés

1.1 Historique

Le premier marché d'options moderne est apparu en 1973 — c'était le CBOE (Chicago Board of Options Exchange). Remarquons (sans prétendre qu'il y a un lien entre les deux évènements) que le célèbre article de Fisher Black et Myron Scholes "The pricing of options and corporate liabilities" a été publié dans la même année. De l'autre côté de l'Atlantique, le premier marché d'options a été créé à Londres en 1978 (London Traded Option Exchange aujourd'hui devenu LIFFE qui fait maintenant partie du groupe Euronext). En France, le MONEP (Marché d'Options NEgotiables de Paris) a ouvert ces portes en 1987.

1.2 Fonctionnement des marchés d'options

Le rôle des marchés organisés d'options est tout d'abord d'assurer la liquidité des options, c'est-à-dire, la cotation permanente et l'existence d'un prix unique pour une option donnée. Deuxièmement, l'existence de marchés organisés simplifie la gestion des contrats optionnels en éliminant le risque de contrepartie, car la contrepartie unique de tous les contrats sur un marché organisé est la chambre de compensation de la bourse. Les différents participants des marchés d'options sont

- Les *brokers* (courtiers) qui exécutent les ordres des investisseurs;
- Les *market makers* (teneurs de marché) qui risquent leur propre capital et assurent la liquidité du marché. Les *spécialistes* affichent en permanence les prix sur les options les plus liquides. Les *contrepartistes* répondent aux demandes de prix.
- La *chambre de compensation* qui sert à éliminer le risque de contrepartie. Elle peut demander (et demande) aux vendeurs d'options le versement des *appels de marge*.

1.3 Différents types de sous-jacents

Les options sont écrites sur

- Actions (CBOE: 1332 actions, MONEP: 67 actions)
- Indices (CAC40, S&P500, indices sectoriels...)
- Taux de change (PHLX)
- Obligations ou swaps
- Futures sur marchandises (commodity): cacao, café, sucre, blé mais aussi des matières premières, l'électricité etc.

.SPX (CBOE) **1199.21 +0.53**

Dec 14, 2004 @ 12:15 ET (Data 15 Minutes Delayed)

Calls	Last Sale	Net	Bid	Ask	Vol	Open Int	Puts	Last Sale	Net	Bid	Ask	Vol	Open Int
04 Dec 1190. (SPT LR-E)	12.00	-1.00	11.20	12.80	897	29000	04 Dec 1190. (SPT XR-E)	2.25	-0.60	2.10	2.60	1579	20741
04 Dec 1195. (SPT LS-E)	8.30	-1.20	7.80	9.00	1882	12938	04 Dec 1195. (SPT XS-E)	4.00	-0.50	3.50	4.30	1782	5846
04 Dec 1200. (SZP LT-E)	5.60	-0.70	5.10	6.00	12635	56793	04 Dec 1200. (SZP XT-E)	6.20	-0.20	5.70	6.60	8840	24317
04 Dec 1205. (SZP LA-E)	3.50	-0.70	3.10	3.80	2404	12758	04 Dec 1205. (SZP XA-E)	8.90	-2.10	8.30	9.50	2	100
05 Jan 1190. (SPT AR-E)	25.00	pc	23.60	25.60	0	418	05 Jan 1190. (SPT MR-E)	14.00	+0.20	13.50	14.80	210	2408
05 Jan 1195. (SPT AS-E)	21.20	+2.60	20.60	22.60	32	647	05 Jan 1195. (SPT MS-E)	15.50	-0.90	15.20	16.80	7	1687
05 Jan 1200. (SZP AT-E)	18.80	-0.20	18.20	18.80	2296	32656	05 Jan 1200. (SZP MT-E)	18.00	--	17.30	18.50	3279	7301
05 Jan 1205. (SZP AA-E)	15.40	-1.10	15.40	17.00	52	0	05 Jan 1205. (SZP MA-E)	20.20	-1.50	19.70	21.30	9	2

Figure 1.1: Exemple de cotation (les options les plus liquides sur l'indice S&P 500). Ask = le prix auquel vous pouvez acheter l'option en question. Bid = le prix auquel vous pouvez le vendre.

Les marchés les plus liquides sont ceux d'options sur indices. Ce sont aussi des marchés où on a le plus besoin d'une calibration précise (car les fourchettes bid–ask sont étroites) et où on a le maximum de données disponibles (par exemple, sur CBOE on trouve plus de 500 options de 8 maturités différentes allant de quelques jours à deux ans sur l'indice S&P 500).

1.4 Les options européennes

Les options européennes et américaines sont des options les plus liquides, leurs prix sont connus et peuvent donc être utilisés comme entrées des algorithmes de calibration. Les options sur indices sont typiquement européennes.

Pour les options européennes l'exercice est possible uniquement à la maturité T; le payoff ne dépend que du prix de sous-jacent à la date finale.

$$\text{Call:} \quad H_T = (S_T - K)^+ \qquad \text{Put:} \quad H_T = (K - S_T)^+$$

Terminologie

- K : *strike* (prix d'exercice).
- Une option est *à la monnaie* à la date t si à cette date $K = S_t$.
- Une option est *dans la monnaie*: si elle devait expirer aujourd'hui son pay-off serait positif.
- La *valeur intrinsèque* d'une option est la quantité d'argent qu'elle rapporterait si elle devait expirer aujourd'hui (la valeur du pay-off aujourd'hui).
- La *valeur temps* d'une option est égale à son prix moins sa valeur intrinsèque.

Les options les plus liquides sont les options à la monnaie (ATM) et, dans une moindre mesure, les options hors de la monnaie (OTM).

Un exemple de feuille de prix d'options européennes est donne sur fig. 1.1.

Propriétés des prix de calls/puts Toutes les propriétés ci-dessous découlent de l'absence d'opportunités d'arbitrage sur les marchés d'options (si on suppose que le prix d'achat coïncide avec le prix de vente).

- Parité call-put
$$\text{Call}(T,K) - \text{Put}(T,K) = S_t - KB(t,T),$$
où $B(t,T)$ est le prix à l'instant t d'un zéro-coupon de maturité T (une obligation qui paie une unité d'argent à la maturité et ne verse pas de coupon). Si la courbe des taux est plate au niveau r, $B(t,T) = e^{-r(T-t)}$. En conséquence de la parité call-put, on obtient les bornes suivantes pour les prix de calls / puts:
$$(S_t - K)^+ < (S_t - KB(t,T))^+ \leq \text{Call}(T,K) \leq S_t. \tag{1.1}$$
$$(KB(t,T) - S_t)^+ \leq \text{Put}(T,K) \leq KB(t,T). \tag{1.2}$$
La valeur temps d'une option Call sur un actif ne versant pas de dividende est donc toujours positive.

- Le prix d'une option call est décroissant par rapport au strike (et le prix d'une put est croissant)
$$K_1 \leq K_2 \quad \Rightarrow \quad \text{Call}(T,K_1) \geq \text{Call}(T,K_2).$$
Cette propriété découle de l'existence de la stratégie call spread qui consiste à acheter une call de strike K_1 et à vendre une call de strike K_2. Comme cette stratégie a un pay-off positif, son prix doit être positif.

- Les prix des calls/puts sont convexes par rapport au strike. Cette propriété correspond à la stratégie butterfly qui consiste à acheter une call de strike K_1, vendre deux calls de strike $(K_1 + K_2)/2$ et acheter une call de strike K_2. On vérifie que cette stratégie a également un pay-off positif dans tous les états de la nature ce qui implique la convexité.

- Le prix d'une option est croissant avec la maturité: $T_1 \leq T_2$ implique $\text{Call}(T_1,K) \leq \text{Call}(T_2,K)$. Cette propriété correspond à la stratégie calendar spread: acheter une call de maturité T_2 et vendre une call de même strike de maturité T_1.

Cas d'un sous-jacent versant des dividendes Pour des options sur un sous-jacent versant des dividendes la relation de parité call-put est modifiée. A la maturité nous avons toujours
$$\text{Call}_T(T,K) - \text{Put}_T(T,K) = S_T - K,$$
cependant pour percevoir ce flux à la maturité, il n'est pas nécessaire d'investir $S_t - B(t,T)K$ à la date t. Si l'action verse des dividendes discrètes connues $D_1, \ldots, D_N$ aux dates $t_1, \ldots, t_n$, alors en achetant une action et en empruntant
$$\sum_{i=1}^{N} D_i B(t,t_i) + KB(t,T)$$
à la banque à la date t, on aura le flux $S_T - K$ à la date T. La parité call-put devient donc
$$\text{Call}_t(T,K) - \text{Put}_t(T,K) = S_t - \sum_{i=1}^{N} D_i B(t,t_i) - KB(t,T).$$
Pour les indices contenant plusieurs actions, on utilise un général l'approximation de taux de dividende continu, i.e., on suppose que l'indice S_t verse en continu une dividende égale à $qS_t dt$. Dans ce cas il est facile de voir que pour s'assurer le flux S_T à la maturité, on doit investir le montant $S_t D(t,T)$ à la date t, où $D(t,T) = e^{-q(T-t)}$. La parité call-put devient donc
$$\text{Call}_t(T,K) - \text{Put}_t(T,K) = S_t D(t,T) - KB(t,T).$$
La dérivation des bornes analogues à (1.1)–(1.2) est laissée en exercice.

1.5 Les options américaines

Pour les options américaines, l'exercice est possible à toute date t avant la maturité T ou à la maturité. Le prix d'une option américaine est donc en général supérieur au prix de l'option européenne correspondante. La différence entre les deux prix s'appelle la *prime d'exercice anticipée*. Dans le cas particulier de l'option call américaine sur un actif ne versant pas de dividende, comme la valeur temps est toujours positive, il n'est jamais optimal d'exercer l'option avant la maturité, et donc, le prix de l'option américaine est égal au prix de l'option européenne.

1.6 Les options exotiques

- Options à barrière: le paiement a lieu (n'a pas lieu) si le sous-jacent a dépassé un niveau contractuel (la barrière) avant cette date. Exemple (up and out call)

$$H_T = (S_T - K)^+ 1_{M_T<B}, \quad \text{où} \quad M_T = \max_{t\le u\le T} S_u$$

  - Moins chère que la call standard

- Options asiatiques: le payoff dépend de la valeur moyenne de sous-jacent pendant la vie de l'option (pour empêcher la manipulation des prix):

$$H_T = \left(\frac{1}{T-t}\int_t^T S_u du - K\right)^+$$

- Options multi-sousjacent: sur un panier d'actions, un panier de taux de change etc.

$$H_T = \left(\sum_{i=1}^n w_i S_T^i - K\right)^+$$

- Les options forward start. Le strike d'une telle option est déterminé à une date future selon une règle spécifié, par exemple, le pay-off peut être

$$H_T = (S_T - mS_{T_0})^+,$$

  où $T_0 < T$ est une date future et m est un nombre fixé dans le contrat (moneyness de l'option).

- Les options cliquets. Une option cliquet peut être vu comme une séquence d'options forward start, avec des dates de départ de plus en plus éloignées. Par exemple le pay-off d'une telle option de maturité $T = 1$ an peut être donné par

$$\sum_{i=1}^{12} \max(0, r_i), \quad r_i = S_{\frac{i}{12}} - S_{\frac{i-1}{12}}.$$

  Cette option permet alors de récupérer l'ensemble des performances positives d'un indice de référence sans en subir des performances négatives.

Partie I

Modèles

Chapitre 2

Le modèle de Black et Scholes et la couverture en delta

2.1 Le modèle de Black et Scholes

Les hypothèses du modèle de Black-Scholes:

- Marchés efficients:
  - Pas de coûts de transaction
  - Pas de restrictions sur le volume de transactions
  - Pas d'opportunité d'arbitrage
- Les rendements du sous-jacent sont gaussiens, stationnaires et indépendants:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t$$

- Le placement à la banque est sans risque et le taux d'intérêt r est constant.

2.1.1 Portefeuille autofinançant

Définition 1. Une stratégie autofinançante est une stratégie dont la valeur n'est pas modifié par l'ajout où retrait du cash.

Soit δ_t le nombre d'actions dans le portefeuille, B_t le solde du compte associé au portefeuille et V_t la valeur du portefeuille à l'instant t: $V_t = \delta_t S_t + B_t$. En temps discret (e.g., si le trading à lieu uniquement aux dates $(t_i)_{i=1}^N$), la condition d'autofinancement s'écrit:

$$\begin{aligned} & B_{t_i-} + \delta_{t_i-} S_{t_i} = B_{t_i} + \delta_{t_i} S_{t_i} \\ \Rightarrow \quad & e^{r(t_i - t_{i-1})} B_{t_{i-1}} + \delta_{t_{i-1}} S_{t_i} = B_{t_i} + \delta_{t_i} S_{t_i} \\ V_{t_i} - V_{t_{i-1}} &= (e^{r(t_i - t_{i-1})} - 1) B_{t_{i-1}} + \delta_{t_{i-1}} (S_{t_i} - S_{t_{i-1}}) \\ &= (e^{r(t_i - t_{i-1})} - 1)(V_{t_{i-1}} - \delta_{t_{i-1}} S_{t_{i-1}}) + \delta_{t_{i-1}} (S_{t_i} - S_{t_{i-1}}) \end{aligned}$$

En passant à la limite $\Delta t \to 0$, on trouve

$$dV_t = (V_t - \delta_t S_t) r dt + \delta_t dS_t$$

Supposons que $V_t = C(t, S_t)$ pour une fonction régulière $C(t, S)$. Alors $dV_t = dC(t, S_t)$ et la formule d'Itô implique:

$$\begin{aligned} dC(t,S_t) &= \frac{\partial C}{\partial t}dt + \frac{\partial C}{\partial S}dS_t + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}d\langle S\rangle_t \\ &= \sigma S_t \frac{\partial C}{\partial S}dW_t + \left(\frac{\partial C}{\partial t} + \mu S_t \frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 C}{\partial S^2}\right)dt \end{aligned}$$

Pour que $dV_t = dC(t, S_t)$, nécessairement

$$\begin{aligned} \delta_t &= \frac{\partial C(t,S_t)}{\partial S} \\ rC(t,S) &= \frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} \end{aligned} \tag{2.1}$$

Si une fonction quelconque $C(t, S)$ satisfait (2.1), le portefeuille avec $\delta_t = \frac{\partial C}{\partial S}$ et $B_t = C(t, S_t) - \delta_t S_t$ est autofinançant. Par absence d'arbitrage, ceci implique que le prix d'une option de payoff $h(S_T)$ est donné par l'unique solution de (2.1) avec condition finale $C(T, S) = h(S)$.

2.1.2 Evaluation risque-neutre

Sous la probabilité historique P,

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t.$$

Soit Q une probabilité équivalente à P définie par:

$$\frac{dQ}{dP}\Big|_{\mathcal{F}_t} = \exp\left(-\frac{\mu - r}{\sigma}W_t - \frac{(\mu - r)^2}{2\sigma^2}t\right)$$

Alors par le théorème de Girsanov $\hat{W}_t := W_t + \frac{\mu - r}{\sigma}t$ est un mouvement Brownien sous Q et

$$\frac{dS_t}{S_t} = rdt + \sigma d\hat{W}_t.$$

Posons $\tilde{C}(t, S_t) := e^{-r(T-t)}E^Q[h(S_T)|S_t]$. Alors $e^{-rt}\tilde{C}(t, S_t)$ est une Q-martingale. Une application de la formule d'Itô donne

$$e^{rt}d(e^{-rt}\tilde{C}) = \left(-r\tilde{C} + \frac{\partial \tilde{C}}{\partial t} + rS_t\frac{\partial \tilde{C}}{\partial S} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 \tilde{C}}{\partial S^2}\right)dt + \sigma S_t \frac{\partial \tilde{C}}{\partial S}d\hat{W}_t,$$

ce qui montre que $\tilde{C}$ est une solution de (2.1). De l'unicité de solution on déduit que le prix d'une option de payoff $h(S_T)$, solution de (2.1), dans le modèle de Black et Scholes est donné par

$$C(t, S_t) = e^{-r(T-t)}E^Q[h(S_t e^{r(T-t) - \frac{1}{2}\sigma^2(T-t) + \sigma\hat{W}_{T-t}})|S_t] \tag{2.2}$$

Ceci est un cas particulier du célèbre *théorème fondamental de valorisation des actifs financiers:* [38, 21] qui dit que le marché est sans opportunité d'arbitrage si et seulement s'il existe une probabilité Q (probabilité risque-neutre), équivalente à la probabilité historique P tel que les prix actualisés de tous les actifs sont des martingales.

2.1.3 Formule de Black et Scholes

Pour une option Call européenne $h(S_T) = (S_T - K)^+$, et (2.2) devient (le calcul de l'espérance est laissé au lecteur)

$$\begin{aligned} \Rightarrow \quad \text{Call}(t, S_t, T, K) &= e^{-r(T-t)}E^Q[(S_t e^{r(T-t) - \frac{1}{2}\sigma^2(T-t) + \sigma\hat{W}_{T-t}} - K)^+|S_t] \\ &= S_t N(d_1) - Ke^{-r(T-t)}N(d_2) \\ \text{avec} \quad d_1 &= \frac{1}{\sigma\sqrt{T-t}}\log\left(\frac{S_t}{Ke^{-r(T-t)}}\right) + \frac{1}{2}\sigma\sqrt{T-t}, \quad d_2 = d_1 - \sigma\sqrt{T-t}. \end{aligned} \tag{2.3}$$

Pour couvrir une Call Européenne de maturité T et prix d'exercice K, il faut donc

- Acheter $\delta_t = \frac{\partial \text{Call}(t,S_t,T,K)}{\partial S} \equiv N(d_1)$ actions
- Placer $B_t = \text{Call}(t, S_t, T, K) - \delta_t S_t$ à la banque

La formule de Black-Scholes est vraie dans le cadre plus général où S est la valeur d'un portefeuille autofinançant quelconque. Par conséquent, dans la situation où le sous-jacent n'est pas un portefeuille autofinançant (si c'est un taux d'intérêt où une action versant des dividendes), pour établir la formule de pricing d'un call, il faut calculer la valeur V_t du portefeuille autofinançant qui est égal au sous-jacent à l'echéance, et remplacer S_t par V_t partout dans la formule (2.3). La sections suivante développe cette approche dans le cas d'options sur un actif versant des dividendes.

2.1.4 Option sur un actif versant des dividendes

Pour les actions, on utilise en général les prévisions de dividendes calculées par les analystes financiers et publiées par les fournisseurs de données. Pour les indices, étant donné que le nombre d'actions est typiquement élevé et ils paient des dividendes quasiment en continu, on utilise généralement l'approximation de taux de dividende continu, qui consiste à dire que le montant de dividende payé entre t et $t + dt$ est égal à $qS_t dt$. Si Y est un portefeuille autofinançant contenant δ_t actions, alors

$$dY_t = \delta_t dS_t + q\delta_t S_t dt.$$

D'un autre côté, la formule d'Itô donne (puisque δ est croissant et donc a variation finie)

$$dY_t = \delta_t dS_t + S_t d\delta_t.$$

Ceci implique $d\delta_t = q\delta_t dt$ et donc $\delta_t = \delta_0 e^{qt}$: un portefeuille autofinançant qui contient e^{-qt} actions à l'instant $t = 0$, va contenir exactement une action à l'instant t. La formule de Black-Scholes est alors vraie dans sa forme habituelle pour le processus Y_t et si on la réécrit en terme de S_t, elle devient

$$\text{Call}(t, S_t, T, K) = S_t e^{-q(T-t)} N(d_1) - K e^{-r(T-t)} N(d_2) \tag{2.4}$$

$$\text{avec} \quad d_{1,2} = \frac{1}{\sigma\sqrt{T-t}} \log\left(\frac{S_t e^{-q(T-t)}}{K e^{-r(T-t)}}\right) \pm \frac{1}{2}\sigma\sqrt{T-t}, \tag{2.5}$$

autrement dit, nous avons remplacé S_t par $S_t e^{-q(T-t)}$ partout dans la formule.

2.1.5 Exemples de couverture en delta

Figure 2.1 montre deux exemples d'évolution de portefeuille de couverture et de l'option qu'on cherche à répliquer dans le modèle de Black-Scholes avec rebalancement très fréquent (8 fois par jour) du portefeuille de couverture. De façon générale, à la fin de chaque trajectoire, le delta est égale soit à 0 si l'option termine hors de la monnaie soit à 1 si l'option termine dans la monnaie.

Figure 2.2 montre l'effet d'un fréquence de rebalancement moins élevée: à gauche, le portefeuille est recalculé une fois par jour et à droite la couverture reste statique pendant toute la période (2 mois), ce qui conduit à une grande erreur de réplication.

Enfin, figure 2.3 montre que même en cas de rebalancement quasi-continu, un saut dans le prix ne peut pas être pris en compte par la couverture en delta et conduit à une erreur de réplication.

2.1.6 Analyse de l'erreur de couverture

Dans cette section nous nous mettons dans le cas $r = 0$ (il est toujours possible de se ramener à cette situation en travaillant avec des quantités actualisées). Nous supposons que le rebalancement de portefeuille de couverture à lieu aux instants ih, $i = 1, 2, \ldots$. Son évolution s'écrit alors

$$dV_t = \frac{\partial C(h[t/h], S_{h[t/h]})}{\partial S} dS_t,$$

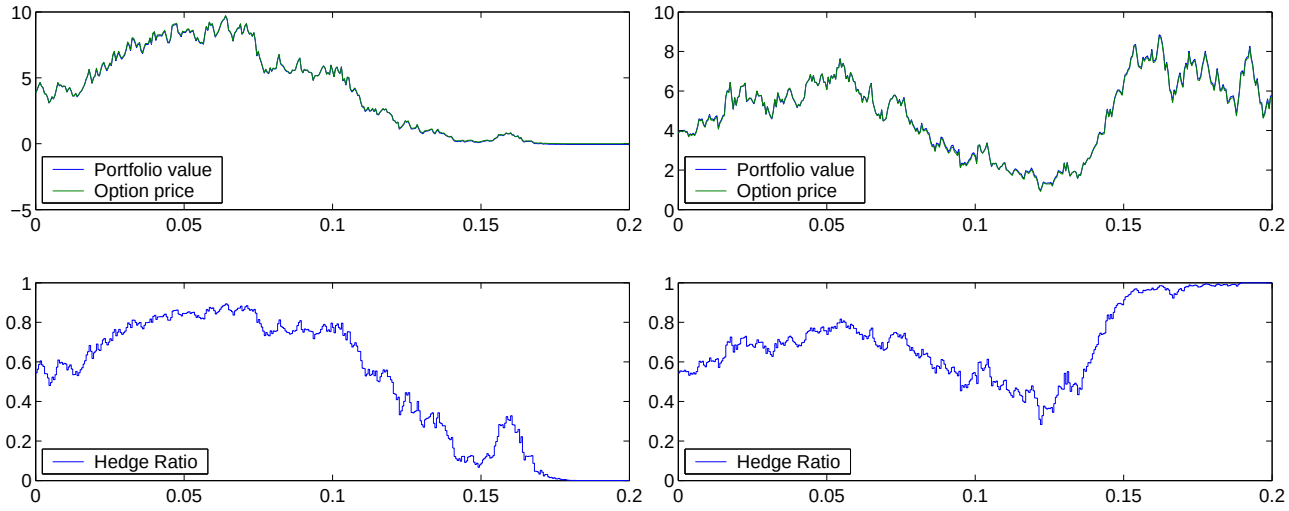


Figure 2.1: Rebalancement continu du portefeuille de couverture (8 fois par jour). Gauche: P&L final $= -0.05$. Droite: P&L final $= 0.08$.

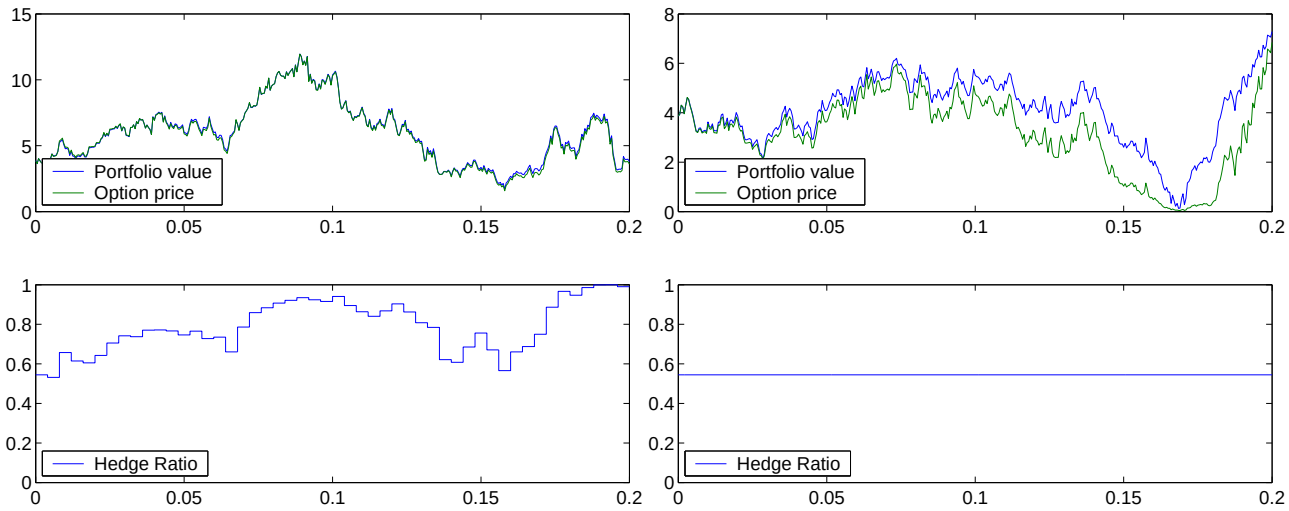


Figure 2.2: Gauche: rebalancement journalier, P&L final $= 0.18$. Droite: couverture statique en delta, P&L final $= 0.44$.

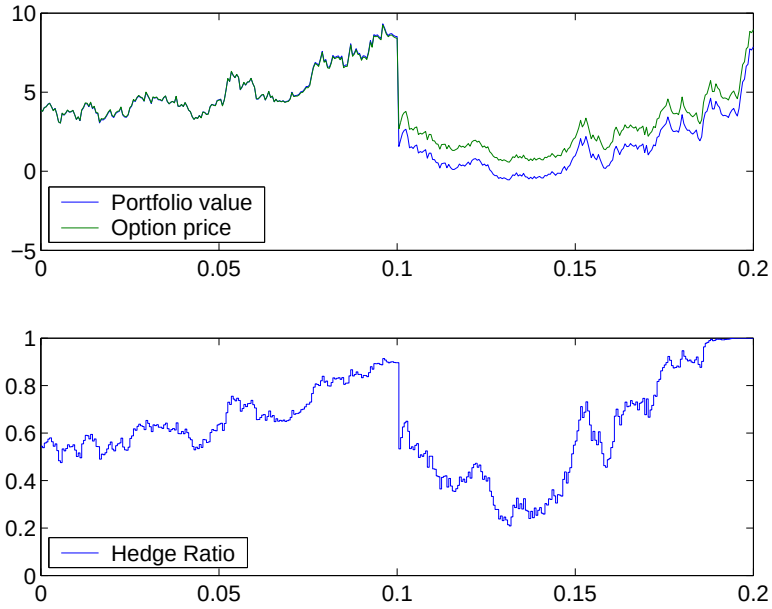


Figure 2.3: Effet d'un saut sur la stratégie de couverture. Rebalancement continu. P&L final = 1.10. Taille du saut : 7%.

où $[x]$ dénote la partie entière de x. D'un autre coté, l'évolution du prix de l'option est

$$dC_t = \frac{\partial C(t, S_t)}{\partial S} dS_t.$$

Pour l'erreur de couverture $\varepsilon_t \equiv C_t - V_t$ nous avons alors

$$d\varepsilon_t = \left(\frac{\partial C(t, S_t)}{\partial S} - \frac{\partial C(h[t/h], S_{h[t/h]})}{\partial S} \right) dS_t.$$

Puisque $\frac{\partial C}{\partial S}$ est borné (dans le cas d'une option erupéenne), (ε_t) est une $\mathbb{Q}$-martingale d'espérance nulle. Introduisons l'erreur de couverture renormalisée $\varepsilon_t^* = \frac{1}{\sqrt{h}} \varepsilon_t$. Il est alors facile de démontrer que $E[(\varepsilon_T^*)^2]$ converge vers une limite bien définie (voir aussi [58]).

Proposition 1. *Supposons que le prix de l'action suit le modèle Black-Scholes avec $r = 0$, et que le pay-off de l'option est tel que $\frac{\partial C}{\partial S}$ est borné. Alors*

$$\lim_{h \to 0} E^Q[(\varepsilon_T^*)^2] = E^Q \left[\frac{1}{2} \int_0^T \left(\frac{\partial^2 C(t, S_t)}{\partial S^2} \right)^2 S_t^4 \sigma^4 dt \right]$$

dès que l'espérance dans la partie droite est finie.

Preuve. Dans cette démonstration, E denote l'espérance sous la probabilité risque-neutre E^Q. Soit $\delta_t = \frac{\partial C(t, S_t)}{\partial S}$, $\phi(t) = h[t/h]$ et $\psi(t) = h + \phi(t)$. En utilisant la bornitude de δ, par isométrie d'Itô on a:

$$E[(\varepsilon_T^*)^2] = \frac{1}{h} E \left[\int_0^T (\delta_t - \delta_{\phi(t)})^2 \sigma^2 S_t^2 dt \right]$$

En utilisant l'intégration par parties entre $\phi(t)$ et $\psi(t)$ puis la formule d'Itô ceci devient

$$\begin{aligned}
E[(\varepsilon_T^*)^2] &= \frac{\sigma^2}{h} E\left[\int_0^T (\psi(t)-t) d\left((\delta_t - \delta_{\phi(t)})^2 S_t^2\right)\right] \\
&= \frac{2\sigma^2}{h} E\left[\int_0^T (\psi(t)-t)(\delta_t - \delta_{\phi(t)}) S_t^2 d\delta_t\right] + \frac{2\sigma^2}{h} E\left[\int_0^T (\psi(t)-t)(\delta_t - \delta_{\phi(t)})^2 S_t dS_t\right] \\
&+ \frac{\sigma^2}{h} E\left[\int_0^T (\psi(t)-t) S_t^4 \left(\frac{\partial^2 C(t,S_t)}{\partial S^2}\right)^2 \sigma^2 dt\right] \\
&+ \frac{\sigma^2}{h} E\left[\int_0^T (\psi(t)-t)(\delta_t - \delta_{\phi(t)})^2 S_t^2 \sigma^2 dt\right] + \frac{4\sigma^2}{h} E\left[\int_0^T (\psi(t)-t)(\delta_t - \delta_{\phi(t)}) S_t^3 \sigma^2 \frac{\partial^2 C(t,S_t)}{\partial S^2} dt\right]
\end{aligned} \tag{2.6}$$

En utilisant le fait que $0 \leq \frac{\psi(t)-t}{h} \leq 1$, que δ_t est borné et $\delta_{\phi(t)} \to \delta_t$ quand $h \to 0$, et le théorème de convergence dominée, nous obtenons immédiatement que les deux termes sur la dernière ligne ci-dessus convergent vers 0 lorsque $h \to 0$. De plus, pour toute fonction intégrable mesurable f,

$$\int_0^T \frac{\psi(t)-t}{h} f(t) dt \to \frac{1}{2} \int_0^T f(t) dt.$$

Avec le théorème de convergence dominée, ceci donne

$$\frac{\sigma^4}{h} E\left[\int_0^T (\psi(t)-t) S_t^4 \left(\frac{\partial^2 C(t,S_t)}{\partial S^2}\right)^2 dt\right] \to \frac{\sigma^4}{2} E\left[\int_0^T S_t^4 \left(\frac{\partial^2 C(t,S_t)}{\partial S^2}\right)^2 dt\right]$$

lorsque $h \to 0$. Il reste à traiter les deux termes de la 2ème ligne de (2.6). Pour le 2ème de ces termes, on a

$$\begin{aligned}
\left|\frac{2\sigma^2}{h} E\left[\int_0^T (\psi(t)-t)(\delta_t - \delta_{\phi(t)})^2 S_t dS_t\right]\right| &\leq \frac{2\sigma^2}{h} E\left[\left(\int_0^T (\psi(t)-t)(\delta_t - \delta_{\phi(t)})^2 S_t dS_t\right)^2\right]^{\frac{1}{2}} \\
&= \frac{2\sigma^2}{h} E\left[\int_0^T (\psi(t)-t)^2 (\delta_t - \delta_{\phi(t)})^4 S_t^4 \sigma^2 dt\right]^{\frac{1}{2}} \\
&\leq 2\sigma^4 E\left[\int_0^T (\delta_t - \delta_{\phi(t)})^4 S_t^4 dt\right]^{\frac{1}{2}} \to 0
\end{aligned}$$

par le théorème de convergence dominée. Passons au 1er terme. En utilisant la formule d'Itô et l'équation de Black-Scholes, on obtient

$$d\delta_t = \frac{\partial^2 C}{\partial S^2} dS_t + \frac{\partial^2 C}{\partial \partial S \partial t} dt + \frac{1}{2} S_t^2 \sigma^2 \frac{\partial^3 C}{\partial S^3} dt = \frac{\partial^2 C}{\partial S^2} dS_t - \frac{\partial^2 C}{\partial S^2} \sigma^2 S_t dt.$$

On a donc

$$\begin{aligned}
&\left|\frac{2\sigma^2}{h} E\left[\int_0^T (\psi(t)-t)(\delta_t - \delta_{\phi(t)}) S_t^2 d\delta_t\right]\right| \\
&\leq \frac{2\sigma^4}{h} E\left[\int_0^T (\psi(t)-t)|\delta_t - \delta_{\phi(t)}| S_t^3 \left|\frac{\partial^2 C}{\partial S^2}\right| dt\right] + \frac{2\sigma^2}{h} E\left[\left(\int_0^T (\psi(t)-t)(\delta_t - \delta_{\phi(t)}) S_t^2 \frac{\partial^2 C}{\partial S^2} dS_t\right)\right]^{\frac{1}{2}} \\
&\leq 2\sigma^4 E\left[\int_0^T |\delta_t - \delta_{\phi(t)}| S_t^3 \left|\frac{\partial^2 C}{\partial S^2}\right| dt\right] + \frac{2\sigma^4}{h} E\left[\int_0^T (\psi(t)-t)^2 (\delta_t - \delta_{\phi(t)})^2 S_t^4 \left(\frac{\partial^2 C}{\partial S^2}\right)^2 dt\right]^{\frac{1}{2}} \to 0,
\end{aligned}$$

en utilisant convergence dominée une dernière fois. □

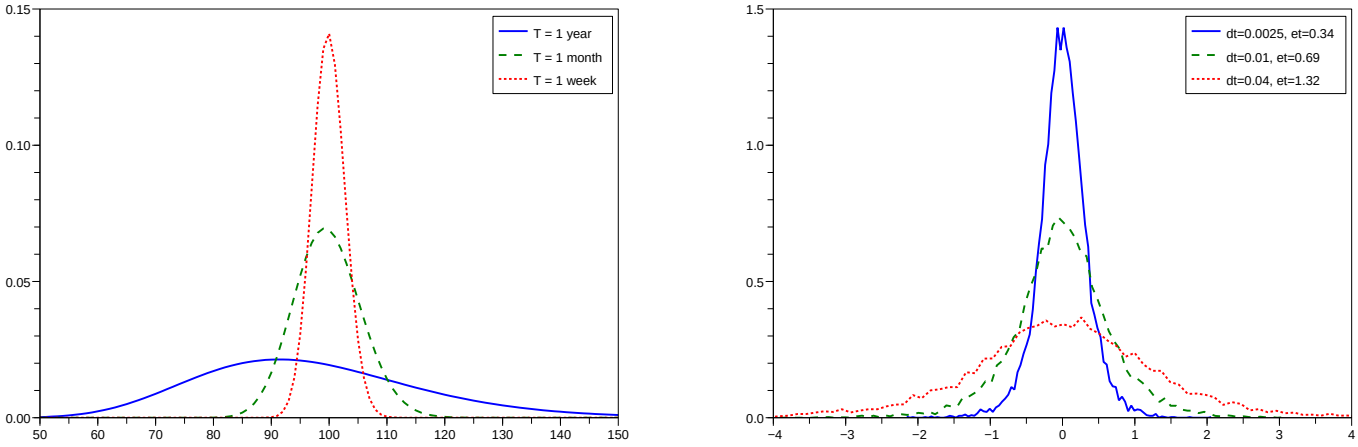


Figure 2.4: Gauche: gamma Black-Scholes d'une option call pour maturités différentes. Droite: histogrammes de l'erreur de couverture pour différentes fréquences de rébalancement.

Exercice 1. En utilisant la forme explicite du gamma Black-Scholes

$$\frac{\partial^2 C(t,S)}{\partial S^2} = \frac{n(d_1)}{S\sigma\sqrt{T-t}},$$

montrer que le résultat précedent s'applique au cas d'une option call.

On peut également démontrer [39] que le processus (ε_t^*) converge en distribution vers un processus bien identifié:

$$(\varepsilon_t^*) \xrightarrow[h\to 0]{d} \frac{1}{\sqrt{2}} \left(\int_0^t \frac{\partial^2 C(u,S_u)}{\partial S^2} \sigma^2 S_u^2 dW'_u \right),$$

où W' est un mouvement Brownien indépendant de W. Ces deux résultats ont plusieurs implications importantes:

- L'erreur de couverture est proportionnelle à la racine du pas de temps $\sqrt{h}$. Pour diviser l'erreur par 2 il faut donc rébalancer le portefeuille 4 fois plus souvent. Ce phénomène est illustré par le graphique de gauche de fig. 2.4.
- L'erreur de couverture est centrée et indépendante de la direction du mouvement du sous-jacent.
- L'erreur de couverture est d'autant plus grande que le gamma de l'option $\frac{\partial^2 C}{\partial S^2}$ est grande. Le gamma Black-Scholes d'une option call est tracée sur le graphique de droite de fig. 2.4 pour des maturités différentes. On voit que les options les plus difficiles à couvrir sont les options ATM proches de l'expiration.

2.1.7 Couverture en gamma

Nous avons vu que pour réduire l'erreur de couverture due à la discrétisation il faut diminuer le gamma du portefeuille. Un moyen de construire un portefeuille gamma-neutre est de rajouter une option ou un autre actif suffisamment liquide dont le gamma compense celui de l'option qu'on cherche à couvrir. Le prix de cet actif de couverture sera noté par C_0.

Supposons que le portefeuille de couverture contient δ_t actions et γ_t actifs C_0, où

$$\gamma_t = \frac{\partial^2 C}{\partial S^2} \Big/ \frac{\partial^2 C_0}{\partial S^2}, \qquad \delta_t = \frac{\partial C}{\partial S} - \frac{\partial C_0}{\partial S} \frac{\partial^2 C}{\partial S^2} \Big/ \frac{\partial^2 C_0}{\partial S^2}.$$

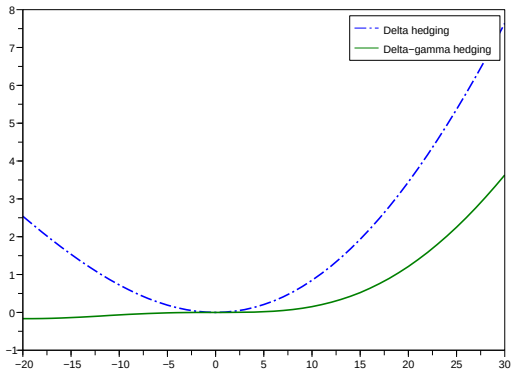


Figure 2.5: Sensibilité des positions couvertes en delta-gamma et en delta aux changements du prix de sous-jacent.

Implication du modèle	**Acceptée par les praticiens?**
– Le risque d'une option peut être complètement éliminé par une position en actions.	Non, car le rébalancement continu est impossible et le modèle est misspecifié
– La quantité d'actions à acheter pour couvrir une option est $\delta_t = \frac{\partial C}{\partial S}$	Oui, c'est la stratégie la plus fondamentale
– Le paramètre fondamental caractérisant une option est sa volatilité σ	Oui, les volatilités des options sont souvent cotées avec leur prix
– La volatilité est la même pour toutes les options sur le même sous-jacent et égale à l'écart type du sous-jacent.	Non, la volatilité est calculée séparément pour chaque option

Table 2.1: Attitude des praticiens vis à vis la formule de Black-Scholes.

Il est facile de voir que le gamma et le delta de ce portefeuille coïncident avec le gamma et le delta de l'option C que l'on cherche à couvrir. Fig. 2.5 montre qu'un portefeuille couvert en gamma-delta reste insensible aux mouvements du sous-jacent d'ampleur beaucoup plus grande que le portefeuille couvert en delta seulement.

2.2 L'utilisation du modèle de Black-Scholes dans les marchés d'options

Le modèle de Black-Scholes reste l'outil fondamental des traders d'options qui se rendent toutefois bien compte des nombreuses inconsistances de ce modèle et de l'impossibilité de l'appliquer tel quel aux marchés d'options. Ce modèle est utilisé comme une approximation de premier ordre au fonctionnement des marchés, et il est ensuite corrigé avec des nombreuses outils et techniques de trading. L'attitude des praticiens vis à vis le modèle de Black-Scholes est grossièrement résumée par le tableau 2.1.

2.2.1 La couverture en delta comme une stratégie de trading

La notion de couverture en delta préconisée par le modèle de Black-Scholes est tellement assimilée par les marchés que l'ensemble d'une option européenne et la couverture en delta associée constitue une stratégie reconnue par les bourses (par exemple, LIFFE propose pas moins de 10 stratégies 'pret-à-porter' delta-neutres). Pour comprendre l'intérêt de la couverture en delta, écrivons le développement limité au premier ordre de la valeur d'un portefeuille couvert en delta (sans supposer le rébalancement continu).

$$C(t+\Delta t, S_{t+\Delta t}) - C(t,S_t) - \frac{\partial C(t,S_t)}{\partial S}\Delta S_t \approx \frac{\partial C(t,S_t)}{\partial t}\Delta t + \frac{1}{2}\frac{\partial^2 C(t,S_t)}{\partial S^2}(\Delta S_t)^2. \tag{2.7}$$

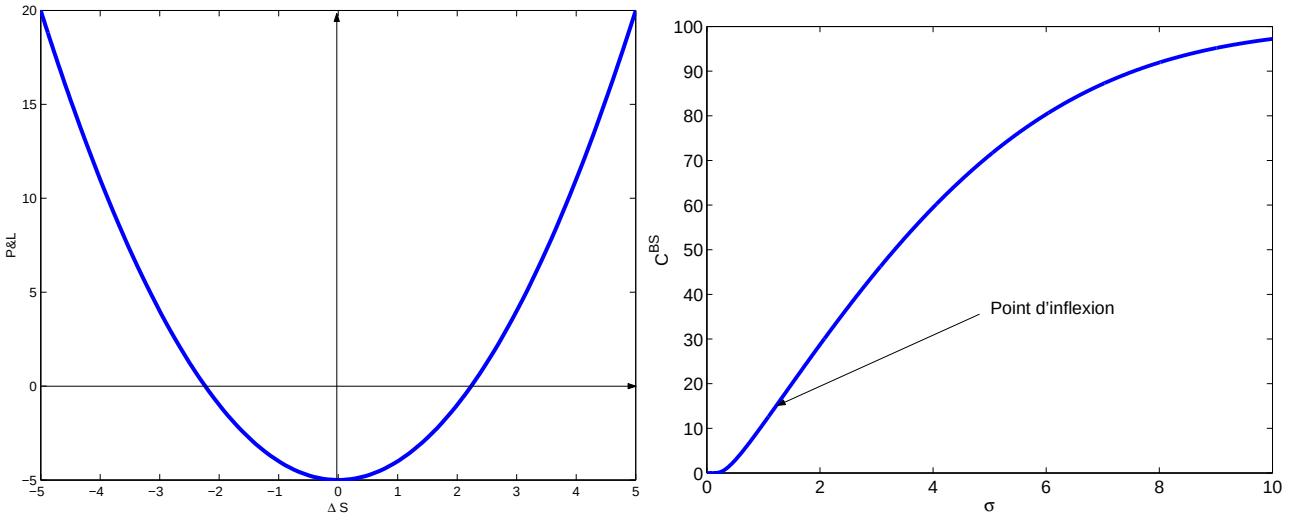


Figure 2.6: Gauche: la variation de la valeur d'une option couvert en delta après un changement du prix de sous-jacent de ΔS. Droite: prix Black-Scholes d'une option en fonction de sa volatilité implicite.

Graphique 2.6 trace la valeur de (2.7) en fonction de ΔS. Il est clair que cette stratégie peut être utilisé par un trader qui anticipe une hausse de volatilité ou un saut de valeur de sous-jacent.

2.2.2 Robustesse de la formule de Black-Scholes

La formule de Black-Scholes est souvent utilisée dans le marché même pour les actifs dont la volatilité n'est pas constante ni même déterministe. Cette pratique est partiellement justifiée par la propriété connue sous le nom de "robustesse de la formule de Black-Scholes": même si la vraie volatilité du sous-jacent est stochastique, la formule de Black-Scholes peut être utilisé pour calculer les bornes supérieures / inférieures du prix, pourvu que la volatilité constante utilisée majore / minore la vraie volatilité.

Supposons que le prix d'action suit

$$\frac{dS_t}{S_t} = \mu dt + \sigma_t dW_t,$$

où σ_t est un processus stochastique, et que l'agent calcule son portefeuille de couverture en utilisant la formule Black-Scholes avec volatilité constante Σ. L'évolution du portefeuille de l'agent est

$$dV_t = rV_t dt + \Delta_{BS}(t, S_t)(dS_t - rS_t dt), \quad V_0 = C_{BS}(0, S_0).$$

L'erreur finale de couverture est donnée par la différence entre V_T et le payoff de l'option:

$$\varepsilon_T = V_T - H_T = V_T - C_{BS}(T, S_T).$$

En appliquant la formule d'Itô on a alors

$$dC_{BS}(t, S_t) = \left(\frac{\partial C_{BS}}{\partial t} + \frac{1}{2} \frac{\partial^2 C_{BS}}{\partial S^2} \sigma_t^2 S_t^2 \right) dt + \frac{\partial C_{BS}}{\partial S} dS_t,$$

mais $C_{BS}(t, S)$ en tant qu'une fonction satisfait l'EDP de Black-Scholes:

$$\frac{\partial C_{BS}}{\partial t} + \frac{1}{2} \Sigma^2 S^2 \frac{\partial^2 C_{BS}}{\partial S^2} = r \left(C_{BS} - S \frac{\partial C_{BS}}{\partial S} \right).$$

En rassemblant les trois équations, on trouve

$$\varepsilon_T = V_T - H_T = \frac{1}{2} e^{rT} \int_0^T e^{-rt} \frac{\partial^2 C_{BS}}{\partial S^2} (\Sigma^2 - \sigma_t^2) S_t^2 dt.$$

Si $\Sigma \geq \sigma_t$ p.s., nous avons $V_T \geq H_T$ p.s.: le portefeuille de couverture Black-Scholes domine le prix de l'option et donc, le prix Black-Scholes domine le vrai prix.

2.3 La volatilité implicite

Dans le modèle de Black et Scholes l'unique paramètre inobservable est la volatilité. Le modèle peut donc être calibré à partir d'un seul prix d'option car la fonction de prix Black-Scholes est strictement croissante en la volatilité. On a

$$\lim_{\sigma\downarrow 0} C^{BS}(\sigma) = (S_t - Ke^{-r(T-t)})^+ \qquad \lim_{\sigma\uparrow\infty} C^{BS}(\sigma) = S_t \tag{2.8}$$

$$\frac{\partial C^{BS}}{\partial \sigma} = Sn(d_1)\sqrt{T-t} \tag{2.9}$$

$$\frac{\partial^2 C^{BS}}{\partial \sigma^2} = \frac{Sn(d_1)\sqrt{T-t}}{\sigma}\left\{\frac{\log^2 m}{\sigma^2(T-t)} - \frac{\sigma^2(T-t)}{4}\right\}, \tag{2.10}$$

où $m = \frac{S_t}{Ke^{-r(T-t)}}$ est le moneyness de l'option. L'équation (2.10) montre que la fonction $\sigma \mapsto C^{BS}(\sigma)$ est convexe sur l'intervalle $(0, \sqrt{\frac{2|\log m|}{T-t}})$ et concave sur $(\sqrt{\frac{2|\log m|}{T-t}}, \infty)$. Ceci implique que l'équation $C^{BS}(\sigma) = C$ pour

$$(S_t - Ke^{-r(T-t)})^+ < C < S_t$$

peut être résolu par la méthode de Newton en utilisant l'algorithme suivant:

$$\sigma_0 = \sqrt{\frac{2|\log m|}{T-t}}$$
$$\sigma_{n+1} = \sigma_n + \frac{C - C^{BS}(\sigma_n)}{\frac{\partial C^{BS}}{\partial \sigma}(\sigma_n)}$$

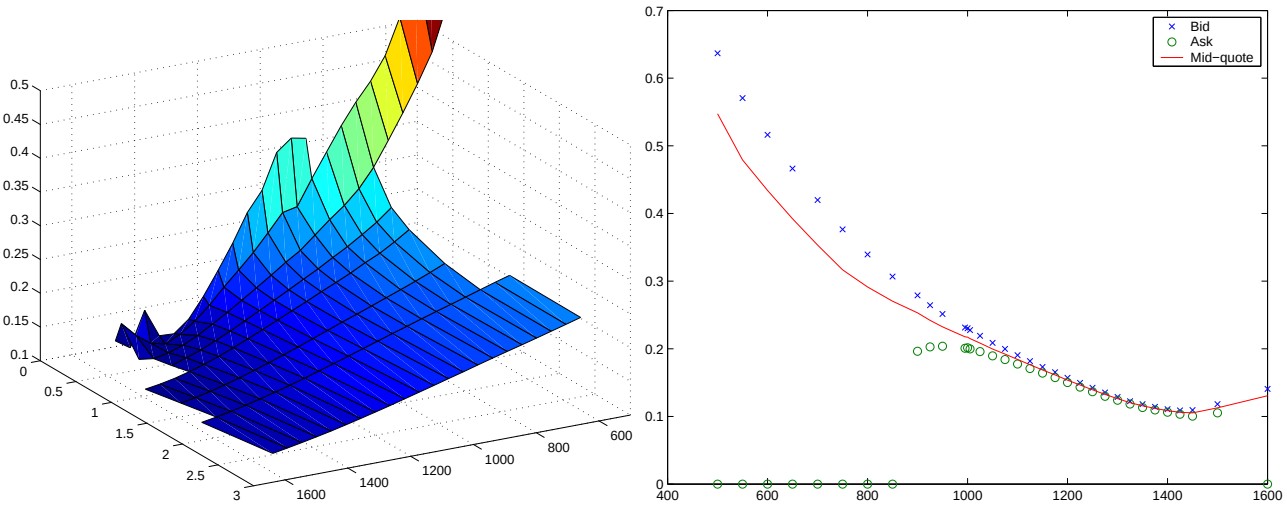
car la série $(\sigma_n) \geq 0$ sera monotone. En pratique, lorsque C est trop proche des bornes d'arbitrage, la dérivée $\frac{\partial C^{BS}}{\partial \sigma}(\sigma_n)$ devient trop petite, ce qui peut conduire à des instabilités numériques. Dans ce cas il est préférable d'utiliser la méthode de bissection. La solution $I(C)$ de l'équation $C^{BS}(\sigma) = C$, où C est le prix d'une option européenne observé sur le marché s'appelle la *volatilité implicite* de cette option. Le modèle de Black-Scholes implique que la volatilité implicite de toutes les options sur le même sous-jacent doit être la même, et égale à la volatilité historique (écart type des rendements annualisé) du sous-jacent. Cependant, lorsqu'on calcule I à partir de prix de différentes options observés sur le marché, on constate que

- La volatilité implicite est toujours supérieure à la volatilité du sous-jacent.
- Les volatilités implicites de différentes options sur le même sous-jacent dépendent de leur strikes et maturités.

Le graphique gauche du fig. 2.3 trace les volatilités implicites des options sur l'indice S&P 500 en fonction de leur strike et maturité observés le 23 janvier 2006. On constate que

- Pour presque tous le strikes, la volatilité implicite décroît en fonction de strike (phénomène de *skew*).
- Pour des très grands strikes on observe parfois une légère remontée de la volatilité implicite (phénomène du *smile*).
- Les phénomènes de smile et skew sont le plus prononcés pour les options de courte maturité; la courbe de volatilité implicite en fonction de strike s'aplatit pour les grandes maturités.

La différence entre la volatilité implicite et la volatilité historique du sous-jacent peut s'expliquer par le fait que la couverture d'une option est plus chère en réalité que dans le modèle de Black-Scholes, à cause, en particulier, des frais de transaction et de la nécessité de couvrir les sources de risque non pris en compte par ce modèle (e.g. le risque de volatilité). Le phénomène de skew est dû au fait que le modèle de Black-Scholes sous-estime la probabilité d'un krach boursier ou d'un grand mouvement de prix en général. Le traders corrigent cette probabilité en augmentant les volatilités implicites des options loin de la monnaie. Finalement, le smile peut être expliqué par les primes de liquidité qui sont plus élevées pour les options loin de la monnaie. Le graphique droit du fig. 2.3 montre que les volatilités implicites des options loin de la monnaie sont presque exclusivement expliquées par les prix Bid (d'achat) qui ont des primes plus élevés pour ces options à cause d'une offre moins importante.



2.3.1 Rôle de la volatilité implicite dans les marchés d'options

La volatilité implicite est très utilisée pour le calcul des ratios de couverture des options européennes. Pour toute option, indépendamment du modèle utilisé on peut écrire:

$$C(t, S_t) = C^{BS}(t, S_t, I),$$

où I est la volatilité implicite de cette option. Si la volatilité implicite ne dépend pas du S_t mais seulement du strike et du temps, on pourrait écrire

$$\frac{\partial C(t, S_t)}{\partial S} = \frac{\partial C^{BS}(t, S_t)}{\partial S},$$

le delta de l'option est donc égal à son delta Black-Scholes. L'absence de dépendance de la volatilité implicite par rapport au sous-jacent a été appelée *le régime sticky strike* par Derman [22]. En réalité cette condition est rarement vérifié, et le changement de I avec le sous-jacent doit être pris en compte:

$$\frac{\partial C(t, S_t)}{\partial S} = \frac{\partial C^{BS}(t, S_t)}{\partial S} + \frac{\partial C^{BS}}{\partial I}\frac{\partial I}{\partial S}. \tag{2.11}$$

Cette divergence entre le delta d'une option et son delta Black-Scholes est souvent traité comme un nouveau source de risque, *risque de vega*[1]. Il est souvent géré en rendant le portefeuille *vega-neutre*. Pour un portefeuille d'options européennes ceci signifie que la somme des dérivées de chaque option par rapport à *sa* volatilité implicite doit être nulle $\sum_{k=1}^{N} \frac{\partial C^{BS}}{\partial \sigma}|_{\sigma=I_k} = 0$. Si maintenant on suppose que les volatilités implicites des différents strikes et maturités varient de la même façon avec le sous-jacent — $\frac{\partial I}{\partial S}$ ne dépend pas de K – le risque dû au changement de volatilité implicite sera totalement couvert.

Une autre approche pour la prise en compte du deuxième terme dans (2.11) consiste à modifier le delta d'une option en faisant des hypothèses appropriés sur l'évolution du smile avec le sous-jacent. Par exemple, dans le régime *sticky delta* ou *sticky moneyness* de Derman, on suppose que la volatilité implicite dépend du ratio K/S mais pas de ces deux variables séparément. En posant $I = I(K/S)$, on a alors

$$\frac{\partial C}{\partial S} = \frac{\partial C^{BS}}{\partial S} - \frac{\partial C^{BS}}{\partial I}\frac{K}{S^2}I'.$$

Comme en général la volatilité décroît avec le strike (pour les options proches de la monnaie), l'hypothèse de sticky delta implique $\frac{\partial C}{\partial S} > \frac{\partial C^{BS}}{\partial S}$: le delta d'une option est supérieur à son delta Black-Scholes.

[1] Vega n'est pas une lettre grecque mais simplement un joli nom pour la dérivée du prix Black-Scholes par rapport à la volatilité implicite ou plus généralement la dérivée du prix d'une option par rapport à la volatilité $\frac{\partial C}{\partial \sigma}$

Exercice 1. On suppose que le profil (smile) de volatilité implicite pour les options près de la monnaie est décroissant avec le strike, et qu'à la monnaie, $\frac{\partial I}{\partial K} = -0.002$. Calculer $\frac{\partial C}{\partial S}$ pour une option Call européenne à la monnaie dans l'hypothèse sticky delta. Les valeurs de paramètres sont $S_0 = 100$, $\sigma = 0.2$, $T = 1$, $r = 0$. Pour le calcul numérique on peut se servir de l'approximation suivante, valable lorsque S_0 est proche de Ke^{-rT}:

$$\mathrm{Call}_{BS}(S_0, K, T) \approx S_0 \left(\frac{1}{2} + \frac{\sigma\sqrt{T}}{5} \right) - Ke^{-rT} \left(\frac{1}{2} - \frac{\sigma\sqrt{T}}{5} \right).$$

Combien d'actions faut-il acheter ou vendre pour couvrir 100 options?

2.3.2 Prise en compte des dividendes dans le calcul de la volatilité implicite

Si on utilise la formule de Black-Scholes (2.3) pour le calcul de la volatilité implicite, pour certains indices (e.g. le DAX) on obtient des résultats raisonnables, mais pour d'autres (CAC 40 ou S&P 500) on obtient des valeurs aberrantes (trop petites), surtout pour des grandes maturités. Ceci est dû à la différence des méthodes de calcul des indices: le DAX est un indice de performance, il peut être vu comme un portefeuille autofinançant d'actions dans lequel les dividendes payés par les actions sont immédiatement réinvestis, ou encore comme une grosse action ne payant pas de dividende. En revanche, CAC 40, S&P 500 etc. sont les indices dits composites, c'est-à-dire, leur valeur est tout le temps égale à la moyenne pondérée des prix des constituants. Si les titres constituant un tel indice paient des dividendes, ceci diminue leur valeur et la valeur de l'indice.

La formule (2.4) implique que la relation de parité Call-Put pour les indices comme CAC 40 prend la forme

$$\mathrm{Call}(S_0, T, K) - \mathrm{Put}(S_0, T, K) = S_0 e^{-qT} - Ke^{-rT},$$

ce qui permet d'estimer la valeur de q par une régression linéaire de

$$\log\{\mathrm{Call}_i - \mathrm{Put}_i + K_i e^{-r_i T_i}\}$$

sur T_i. La volatilité implicite peut alors être calculée en inversant la formule (2.4).

2.4 Complément: formule de Garman-Kohlhagen et cotation d'options sur taux de change

Si le sous-jacent de l'option est un taux de change X_t, alors la valeur d'un portefeuille autofinançant qui rapporte X_T à l'écheance de l'option T est égale au prix d'un zéro-coupon étranger converti en monnaie domestique, donc, dans le cas de taux déterministes,

$$V_t = e^{-\int_t^T r_s^f ds} X_t,$$

où r^f est le taux court étranger. La formule de Black-Scholes, qui est alors connue sous le nom de formule de Garman-Kohlhagen [32], prend la forme

$$\mathrm{Call}(t, X_t, T, K) = X_t e^{-\int_t^T r_s^f ds} N(d_1) - Ke^{-\int_t^T r_s ds} N(d_2) \tag{2.12}$$

$$\text{avec} \quad d_{1,2} = \frac{1}{\sigma\sqrt{T-t}} \log\left(\frac{X_t e^{-\int_t^T r_s^f ds}}{Ke^{-\int_t^T r_s ds}} \right) \pm \frac{1}{2}\sigma\sqrt{T-t}, \tag{2.13}$$

La cotation d'options sur taux de change est sensiblement différente de celles sur actions et indices: alors que dans les marchés *equity*, on affiche les prix de calls et de puts en fonction de leur prix d'exercice, pour les options FX on affiche la volatilité implicite à la monnaie (ATM vol) ainsi que la volatilité implicite des montages *butterfly* (qui represente la convexité du smile de volatilité) ou *risk-reversal* (qui represente la pente du smile) en fonction de leur delta (fig. 2.7). L'utilisation du delta, du volatilité implicite, et des montages qui correspondent aux principaux facteurs de risque associés au smile de volatilité permet de rendre l'affichage plus stable dans le temps et plus lisible pour les traders.

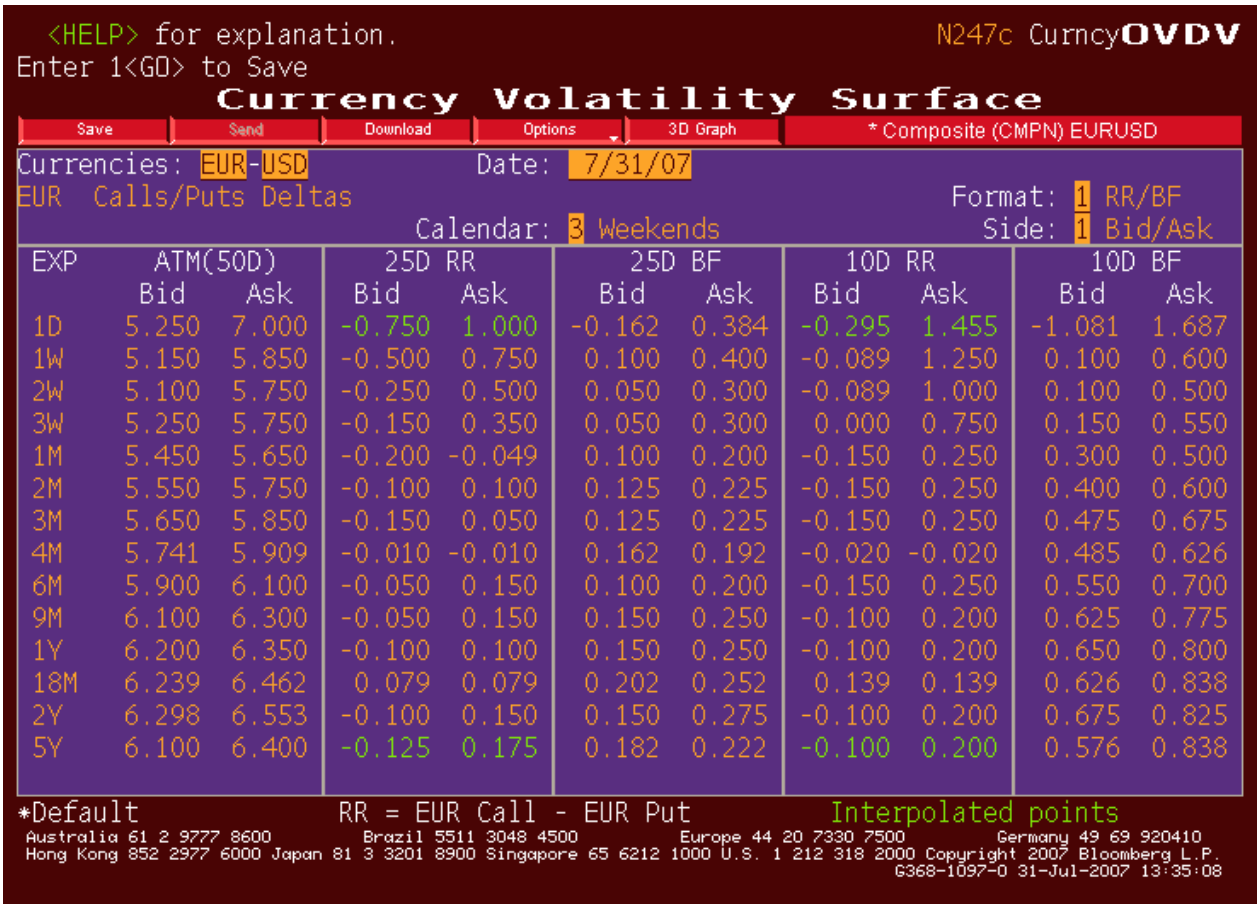


Figure 2.7: Une page Bloomberg avec les cotations des options sur taux de change.

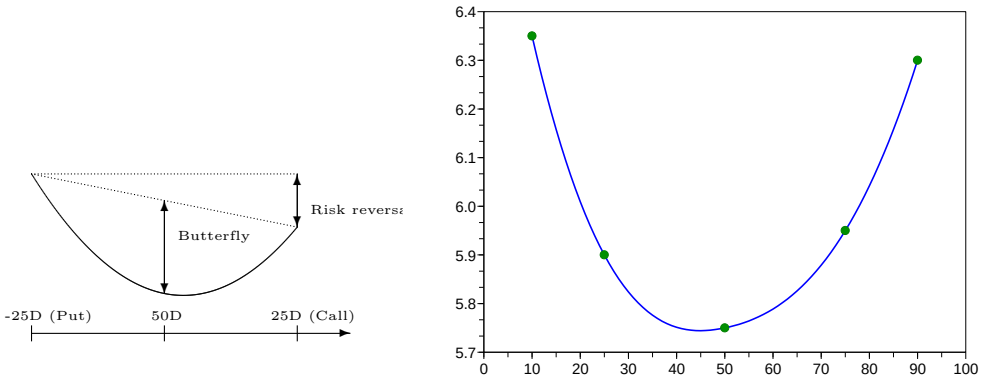


Figure 2.8: A gauche: representation graphique des Butterfly et Risk Reversal. A droite: smile de volatilité pour les options sur EUR/USD de maturité 3 mois.

Cotation en delta Le delta de l'option dans le modèle de Garman-Kolhagen est

$$e^{-\int_0^T r_t^f\,dt}N(d_1), \quad d_1 = \frac{1}{\sigma\sqrt{T}}\log\left(\frac{X_0 e^{-\int_0^T r_t^f\,dt}}{Ke^{-\int_0^T r_t dt}}\right) + \frac{1}{2}\sigma\sqrt{T},$$

où X_0 est le taux de change courant, mais pour la cotation on utilise $N(d_1)$ pour que le "delta" soit compris exactement entre 0 et 1 pour toutes les maturités. Il y a donc une bijection entre le delta de l'option et son stike. L'option "50Δ" correspond à $N(d_1) = 0.5 \iff d_1 = 0$ et a le strike

$$K_{50\Delta} = X_0 e^{\int_0^T (r_t - r_t^f + \frac{\sigma^2}{2})dt}.$$

De même, pour une option "$x\Delta$" on a

$$K_{x\Delta} = X_0 e^{\int_0^T (r_t - r_t^f + \frac{\sigma^2}{2})dt} e^{-\sigma\sqrt{T}N^{-1}(x)}.$$

Risk reversal et butterfly Les "volatilités" du risk-reversal et du butterfly ne sont pas en fait des volatilités mais simplement les mesures de la pente et de la convexité du smile respectivement. Il sont calculés à partir de volatilités des calls et des puts hors la monnaie avec les formules suivantes:

$$BB_{25\Delta} = \frac{\sigma_{25\Delta c} + \sigma_{-25\Delta p}}{2} - \sigma_{50\Delta}, \qquad RR_{25\Delta} = \sigma_{25\Delta c} - \sigma_{-25\Delta p}$$

Ces formules sont illustrées sur le graphique de gauche de fig. 2.8. Pour reconstituer le smile de volatilité implicite, on peut recalculer les volatilités des options call et put hors la monnaie:

$$\sigma_{25\Delta c} = BB_{25\Delta} + \sigma_{50\Delta} + \frac{RR_{25\Delta}}{2}, \qquad \sigma_{-25\Delta p} = BB_{25\Delta} + \sigma_{50\Delta} - \frac{RR_{25\Delta}}{2},$$

puis les interpoler avec une méthode d'interpolation convenable. Remarquons que le butterfly est toujours positif, mais le risk reversal peut changer de signe sur le FX, contrairement aux marchés d'actions, où il est presque toujours négatif.

Chapitre 3

Les modèles à volatilité locale et la diffusion implicite

3.1 Modèles à volatilité locale

Dans la section 2.3 nous avons vu que le modèle de Black-Scholes à volatilité constante ne peut pas reproduire l'ensemble des prix d'options observés sur le marché pour un sous-jacent donné, car leur volatilité implicite varie en fonction du strike et de la maturité. Pour prendre en compte le smile du marché tout un restant dans le cadre markovien et complet (un seul facteur de risque) une solution naturelle est de modéliser la volatilité comme une fonction déterministe du temps et de la valeur du sous-jacent:

$$\begin{aligned} \frac{dS_t}{S_t} &= \mu_t dt + \sigma(t, S_t) dW_t, \\ \frac{dB_t}{B_t} &= r_t dt. \end{aligned} \tag{3.1}$$

Dans ces équations, le taux d'intérêt r_t peut aussi dépendre de t et de S_t, et la dérive μ_t peut être un processus prévisible général car on va s'en débarrasser grâce à l'évaluation risque neutre. Un exemple de modèle de ce type bien étudié dans la littérature est donné par le modèle CEV (Constant Elasticity of Variance) de [18], qui sera analysé en détail à la fin de cette section. Dans ce modèle, la volatilité est une fonction puissance du niveau de sous-jacent:

$$\frac{dS_t}{S_t} = \mu_t dt + \frac{\sigma_0}{S^{1-\alpha}} dW_t \tag{3.2}$$

Le modèle de Black-Scholes et le modèle gaussien sont des cas particuliers de cette équation avec $\alpha = 1$ et $\alpha = 0$ respectivement.

Par le même argument d'autofinancement que dans le modèle Black-Scholes (voir section 2.1), le prix d'une option qui paie $h(S_T)$ à l'instant T satisfait

$$r_t C(t,S) = \frac{\partial C}{\partial t} + r_t S \frac{\partial C}{\partial S} + \frac{1}{2}\sigma(t,S)^2 S^2 \frac{\partial^2 C}{\partial S^2}, \quad C(T,S) = h(S),$$

et le portefeuille autofinançant de couverture contient $\delta_t = \frac{\partial C(t,S_t)}{\partial S}$ actions et $B_t = C(t,S_t) - \delta_t S_t$ en cash. L'équation de pricing garde alors la même forme que dans le modèle de Black-Scholes mais on ne peut plus en déduire une formule explicite car la volatilité dépend maintenant du sous-jacent.

Par analogie exacte au modèle de Black-Scholes, le modèle de volatilité locale correspond à un marché complet (le seul facteur de risque est le sous-jacent) et l'unique probabilité risque-neutre (si elle existe) est donnée par

$$\frac{dQ}{dP}|_{\mathcal{F}_t} = \exp\left(- \int_0^t \frac{\mu_s - r_s}{\sigma_s} dW_s - \int_0^t \frac{(\mu_s - r_s)^2}{2\sigma_s^2} ds \right). \tag{3.3}$$

Sous cette probabilité,

$$\frac{dS_t}{S_t} = r_t dt + \sigma(t, S_t) dW_t^Q. \tag{3.4}$$

Pour assurer l'existence de la probabilité risque-neutre, il faut alors démontrer que (3.3) est bien une martingale. Supposons qu'il existe un processus de *prime de risque* λ_t tel que $\mu_t = r_t + \lambda_t \sigma_t$ et

$$E[e^{\frac{1}{2}\int_0^T \lambda_s^2 ds}] < \infty.$$

Par exemple, il suffit que λ soit bornée. Alors le critère de Novikov (théorème 8 de l'annexe) garantit que

$$\frac{dQ}{dP}|_{\mathcal{F}_t} = \exp\left(-\int_0^t \lambda_s dW_s - \frac{1}{2}\int_0^t \lambda_s^2 ds\right)$$

est une martingale uniformément intégrable et la probabilité risque-neutre Q est bien définie. Sous cette probabilité,

$$\frac{dS_t}{S_t} = r_t dt + \sigma(t, S_t) d\hat{W}_t \quad \text{et} \quad C(t,S) = E^Q[e^{-\int_t^T r_s ds} h(S_T)|S_t = S]$$

Soulignons bien que cette probabilité risque-neutre est unique, c'est-à-dire, elle est complètement déterminé une fois qu'on a estimé le modèle historique (3.1). Les prix d'options sont donc également complètement déterminés, et il n'est pas possible de calibrer le modèle aux prix d'options cotées sur le marché. Pour cette raison, en pratique, le modèle de volatilité locale n'est pas estimé à partir des données historiques mais utilisé directement dans sa forme risque-neutre (3.4) et calibré aux prix d'options cotées.

Modèle CEV Le modèle CEV (3.2) fournit un exemple simple et non-trivial d'un modèle à volatilité locale. Plaçons-nous sous la probabilité risque-neutre, et supposons que le prix forward du sous-jacent $F_t = e^{r(T-t)} S_t$ suit le modèle CEV:

$$dF_t = \sigma_0 F_t^\alpha dW_t, \quad 0 < \alpha \leq 1 \tag{3.5}$$

La valeur 0 est une barrière *absorbante*, si $F_t = 0$ pour un t, $F_s \equiv 0$ pour tout $s \geq t$. La contrainte $0 < \alpha \leq 1$ est imposée puisqu'on peut demontrer que pour $\alpha > 1$, (F_t) est une martingale locale stricte, c'est-à-dire, pas une vraie martingale, ce qui peut conduire, par exemple, à la violation de la parité Call-Put et d'autres inconvenients. Montrons que pour $0 < \alpha \leq 1$, l'équation (3.5) définit une vraie martingale de carré intégrable sur $[0, T]$ pour tout $T < \infty$. Pour cela il est suffisant de voir que

$$E\left\{[F]_T\right\} = E\left\{\sigma_0^2 \int_0^T F_t^{2\alpha} dt\right\} < \infty. \tag{3.6}$$

Soit $\tau_n = \inf\{t : F_t \geq n\}$. $F_{T\wedge\tau_n}$ est alors de carré integrable, et on a, pour $0 < \alpha \leq 1$,

$$E[F_{\tau_n\wedge T}^2] = \sigma_0^2 E\left[\int_0^{\tau_n\wedge T} F_t^{2\alpha} dt\right] \leq \sigma_0^2 E\left[\int_0^{\tau_n\wedge T} (1 + F_t^2) dt\right] \leq \sigma_0^2 E\left[\int_0^T (1 + F_{t\wedge\tau_n}^2) dt\right]$$

Par le lemme de Gronwall on a alors

$$\sigma_0^2 E\left[\int_0^{\tau_n\wedge T} F_t^{2\alpha} dt\right] = E[F_{\tau_n\wedge T}^2] \leq \sigma_0^2 T e^{\sigma_0^2 T},$$

d'où (3.6) est déduit par convergence monotone.

La forme de la volatilité implicite du modèle CEV est connue grâce à l'approximation asymptotique de Hagan et Woodward [37]:

$$\sigma^{imp}(K,T) = \frac{\sigma_0}{F_m^{1-\alpha}}\left\{1 + \frac{(1-\alpha)(2+\alpha)}{24}\left(\frac{F_0 - K}{F_m}\right)^2 + \frac{(1-\alpha)^2}{24}\frac{\sigma_0^2 T}{F_m^{2-2\alpha}} + \dots\right\}, \quad F_m = \frac{1}{2}(F_0 + K).$$

Au premier ordre, on a donc $\sigma^{imp}(K,T) \approx \frac{\sigma_0}{F_m^{1-\alpha}}$: la volatilité implicite a la même forme que la volatilité locale mais avec une pente à la monnaie 2 fois plus petite.

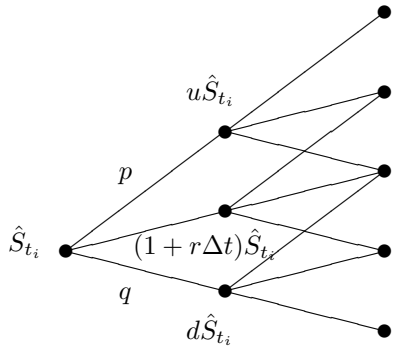


Figure 3.1: Arbre trinomial.

3.2 Arbre trinomial de pricing

Dans cette section, pour simplifier l'exposé, on pose $r_t \equiv r$ pour tout t. Une première méthode pour calculer numériquement les prix des options dans un modèle à volatilité locale consiste à approcher la diffusion en temps continu S par une chaîne de Markov (un arbre) $\hat{S}_{t_i}, i = 0, \ldots, N$, $t_i = t_0 + i\Delta t$, en utilisant un résultat de convergence (théorème 9 de l'annexe) des chaînes de Markov vers les diffusions. Dans un arbre binomial défini par

$$\hat{S}_{t_{i+1}} = \begin{cases} u\hat{S}_{t_i}, & q \\ d\hat{S}_{t_i}, & 1-q \end{cases}$$

la probabilité risque-neutre de transition q est fixée par la condition de martingale: $qu + (1-q)d = 1 + r\Delta t$. Pour tenir compte du fait que la volatilité est variable, on devrait alors faire varier les paramètres u et d en fonction de la position du noeud dans l'arbre [54], mais ceci conduit à des difficultés d'implémentation (difficile par exemple d'obtenir un arbre recombinant). Ce problème ne se pose pas si on utilise un arbre *trinomial*:

$$\hat{S}_{t_{i+1}} = \begin{cases} u\hat{S}_{t_i}, & p \\ (1+r\Delta t)\hat{S}_{t_i}, & 1-p-q, \\ d\hat{S}_{t_i}, & q \end{cases}$$

Pour la convergence vers le processus continu (voir théorème 9) on impose:

$$E[\hat{S}_{t_{i+1}}|\hat{S}_{t_i}] = (1+r\Delta t)\hat{S}_{t_i} \quad \text{(arbre risque-neutre)}$$
$$\mathrm{Var}[\hat{S}_{t_{i+1}}|\hat{S}_{t_i}] = \hat{S}_{t_i}^2\sigma^2(t_i, \hat{S}_{t_i})\Delta t$$

Dans un arbre trinomial on peut donc fixer u et d (le squelette de l'arbre) librement et choisir les probabilités de transition $p(t_i, \hat{S}_{t_i})$ et $q(t_i, \hat{S}_{t_i})$ pour satisfaire les équations ci-dessus. Un calcul rapide laissé au lecteur montre qu'il faut prendre

$$p(t_i, \hat{S}_{t_i}) = \frac{\sigma(t_i, \hat{S}_{t_i})^2\Delta t}{(u-d)(u-1-r\Delta t)} \qquad q(t_i, \hat{S}_{t_i}) = \frac{\sigma(t_i, \hat{S}_{t_i})^2\Delta t}{(u-d)(1+r\Delta t-d)} \tag{3.7}$$

Pour que l'arbre soit recombinant (un arbre est dit recombinant si le nombre de noeuds dans chaque tranche croit linéairement avec temps, comme sur le dessin 3.1, et non pas exponentiellement), on pose $ud = (1+r\Delta t)^2$.

L'arbre n'admet pas d'opportunité d'arbitrage si toutes les probabilités de transition sont positives, et pour cela il suffit que $p + q \leq 1$ dans chaque noeud de l'arbre. Ceci est toujours vrai si le pas de temps Δt est suffisamment petit en comparaison avec le pas d'espace (à comparer avec la condition CFL de stabilité d'un schéma de différences finies explicite). Par exemple, $p + q \leq 1$ si la condition suivante est respectée (les détails du calcul sont laissés au lecteur):

$$1 + r\Delta t - d > \sigma(t_i, \hat{S}_{t_i})\sqrt{\Delta t}.$$

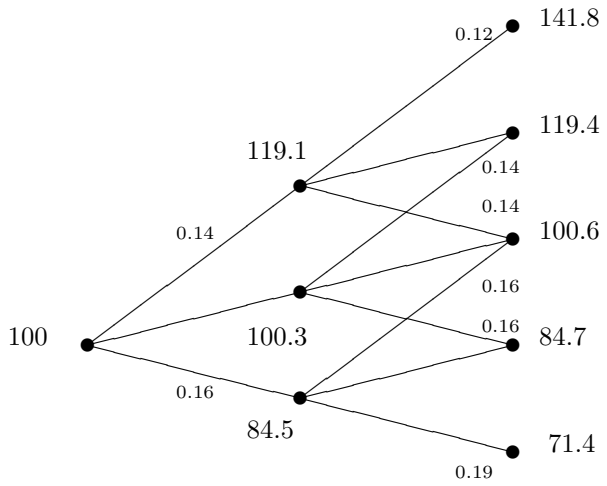


Figure 3.2: Exemple d'un arbre trinomial

Pour que cette inégalité soit toujours vérifiée, on peut poser

$$d = 1 + r\Delta t - \bar{\sigma}\sqrt{\Delta t},$$

où $\bar{\sigma}$ est une constante qui satisfait $\bar{\sigma} > \sigma(t, S)$ dans tous les noeuds de l'arbre. Cette constante n'est pas toujours facile à trouver si $\sigma(t, S)$ n'est pas bornée. Pour contourner cette difficulté, il suffit de choisir $\bar{\sigma}$ assez grand et remplacer le modèle initial par le modèle

$$\frac{dS_t}{S_t} = r_t dt + \min(\bar{\sigma}, \sigma(t, S_t))dW_t.$$

L'arbre construit à partir de cette diffusion sera toujours sans opportunité d'arbitrage, au prix de travailler dans un modèle légèrement différent.

Evaluation backward des prix d'options dans un arbre trinomial Pour calculer le prix d'une option call européen sur un arbre, l'algorithme standard peut être employé:

- A la date finale, les prix sont donnés par: $C(t_N, S) = (S - K)^+$.
- A toute date t_i pour $i < N$, les prix dans tous les noeuds de l'arbre sont calculés à partir des prix à la date t_{i+1} par l'évaluation risque-neutre:

$$C(t_i, S) = \frac{1}{1 + r\Delta t}\{pC(t_{i+1}, uS) + qC(t_{i+1}, dS) + (1 - p - q)C(t_{i+1}, (1 + r\Delta t)S)\}.$$

Exemple 1 (Modèle CEV)**.** Prenons l'exemple du modèle CEV (3.2). Avec les paramètres $S_0 = 100$, $\alpha = 0.5$, $\sigma_0 = 3$, $r = 0.03$, $N = 2$, $\Delta t = 0.1$ et $\bar{\sigma} = 0.5$, on trouve $d \approx 0.845$, et $u \approx 1.191$, ce qui correspond à l'arbre de la figure 3.2. Les prix d'une option call de strike $K = 100$ dans les différents noeuds de l'arbre, obtenus par l'évaluation backward, sont montrés sur fig. 3.3. En itérant la procédure pour plusieurs strikes (avec un arbre plus grand et les données légèrement différentes), on trouve un comportement de volatilité implicite tracé sur fig. 3.4, qui montre que les modèles à volatilité locale reproduisent bien le phénomène de skew de volatilité.

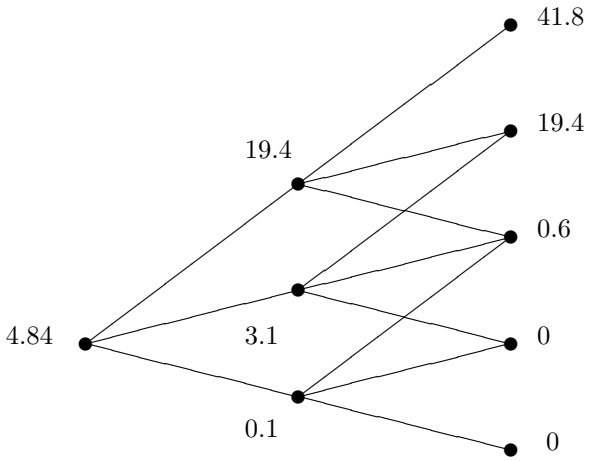


Figure 3.3: Exemple d'évaluation (prix d'une option call dans chaque noeud de l'arbre avec les données de l'exemple 1.)

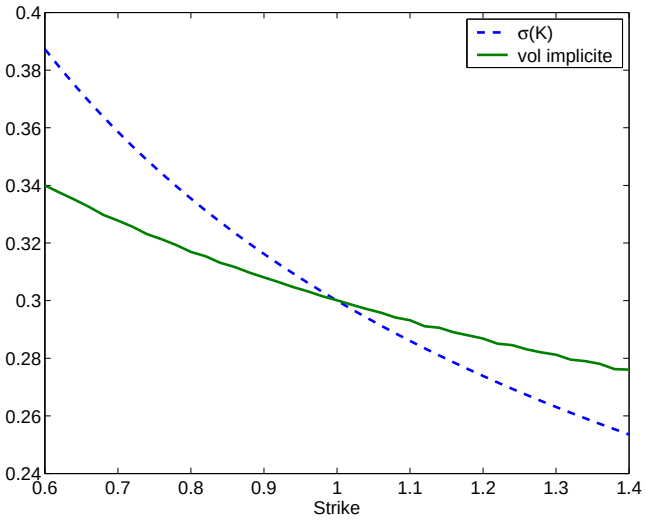


Figure 3.4: Skew (profil décroissant) de volatilité dans le modèle CEV avec $\sigma = 0.3$, $\alpha = 0.5$, $S_0 = 1$ et $T = 1$

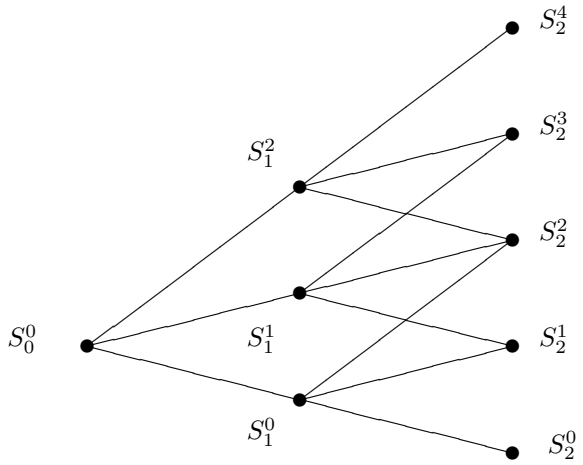


Figure 3.5: Numérotation des noeuds

3.3 Arbres implicites

Dans la section précédente, on a vu comment calculer les prix des options européennes dans un arbre trinomial à partir des probabilités risque-neutres de transition. En réalité, les options européennes sont souvent cotées sur le marché, et au lieu de calculer leurs prix, on les utilise comme données d'entrée pour le pricing des options exotiques. Dans ce cas, on a besoin de construire un arbre trinomial qui soit en accord avec les prix observés sur le marché pour les options européennes.

Un arbre trinomial correspond en général à un marché incomplet (à chaque noeud, il y a trois possibilités d'évolution du sous-jacent, et sur le marché il n'y a que deux actifs). De plus, il existent plusieurs probabilités risque-neutres qui reproduisent les prix d'un ensemble donné d'options européennes. Mais le modèle trinomial récombinant a une propriété qui nous facilite la vie: si on suppose que sur le marché on observe les prix des options call ayant comme strike et maturité les coordonnées t et S de tous les noeuds de l'arbre, alors il existe une seule probabilité risque-neutre, qui reproduit ces prix et *sous laquelle le modèle reste markovien.* On peut expliquer ce phénomène intuitivement comme suit: dans un arbre récombinant markovien, les probabilités de transition ne peuvent dépendre que de la position du noeud dans l'arbre, elles ne peuvent pas dépendre de la trajectoire que le processus du prix a emprunté pour arriver à ce noeud. On a donc un degré de liberté par noeud car il y a deux probabilités p et q et une seule condition (risque-neutralité) pour les déterminer, et la connaissance d'un prix d'option par noeud permet de déterminer cette probabilité de façon unique.

Dans cette section, on verra, comment construire les probabilités de transition dans cet unique arbre trinomial markovien à partir des prix d'options cotées. Cet arbre s'appelle alors l'arbre implicite. Les arbres binomiaux implicites (où on calibre les positions des noeuds) ont été développés par Derman et Kani [23] et Rubinstein [54] et les arbres trinomiaux implicites par Derman et Kani [24].

Arbre trinomial implicite Dans ce qui suit, on notera par S_n^m le m-ième noeud de n-ième génération d'un arbre trinomial (la n-ième génération correspond à l'instant $t_n = n\Delta t$) pour $n = 0, \dots, N$, $m = 0, \dots, 2n$. La numérotation des noeuds est montré sur fig. 3.5.

On a comme données les prix aujourd'hui des calls $C(t_n, S_n^m)$ de maturités t_n, $n = 1, \dots, N$, et pour chaque maturité t_n ayant comme strikes les noeuds $S_n^0, \dots, S_n^{2n}$, et on cherche à calculer les probabilités de transition p_n^m et q_n^m associées à chaque noeud.

La condition de risque-neutralité implique:

$$p_n^m S_{n+1}^{m+2} + (1 - p_n^m - q_n^m) S_{n+1}^{m+1} + q_n^m S_{n+1}^m = (1 + r\Delta t) S_n^m. \tag{3.8}$$

Pour un noeud S_n^m de l'arbre, introduisons l'actif d'Arrow-Debreu, qui paie 1 dollar à l'instant t_n si le processus

de prix se trouve dans le noeud S_n^m. Le prix de cet actif est

$$\lambda_n^m = \frac{1}{(1+r\Delta t)^n} P[\hat{S}_n = S_n^m].$$

En utilisant les actifs d'Arrow-Debreu, on peut écrire le prix d'une call de maturité t_{n+1} et strike S_{n+1}^{m+1} comme

$$\begin{aligned}
C(t_{n+1}, S_{n+1}^{m+1}) &= \frac{1}{(1+r\Delta t)^{n+1}} E[(\hat{S}_{t_{n+1}} - S_{n+1}^{m+1})^+] \\
&= \frac{1}{(1+r\Delta t)^{n+1}} E\{E[(\hat{S}_{t_{n+1}} - S_{n+1}^{m+1})^+ | \hat{S}_n]\} \\
&= \frac{1}{1+r\Delta t} \sum_{j=0}^{2n} \lambda_n^j E[(\hat{S}_{t_{n+1}} - S_{n+1}^{m+1})^+ | \hat{S}_n] \\
&= \frac{1}{1+r\Delta t} \sum_{j=0}^{2n} \lambda_n^j \{p_n^j (S_{n+1}^{j+2} - S_{n+1}^{m+1})^+ + (1 - p_n^j - q_n^j)(S_{n+1}^{j+1} - S_{n+1}^{m+1})^+ \\
&\quad + q_n^j (S_{n+1}^{j} - S_{n+1}^{m+1})^+\} \\
&= \frac{1}{1+r\Delta t} \left\{ \lambda_n^m p_n^m (S_{n+1}^{m+2} - S_{n+1}^{m+1}) + \sum_{j=m+1}^{2n} \lambda_n^j E[\hat{S}_{n+1} - S_{n+1}^{m+1} | \hat{S}_n = S_n^j] \right\} \\
&= \frac{1}{1+r\Delta t} \left\{ \lambda_n^m p_n^m (S_{n+1}^{m+2} - S_{n+1}^{m+1}) + \sum_{j=m+1}^{2n} \lambda_n^j ((1+r\Delta t) S_n^j - S_{n+1}^{m+1}) \right\}
\end{aligned}$$

d'où

$$p_n^m = \frac{(1+r\Delta t) C(t_{n+1}, S_{n+1}^{m+1}) - \sum_{j=m+1}^{2n} \lambda_n^j ((1+r\Delta t) S_n^j - S_{n+1}^{m+1})}{\lambda_n^m (S_{n+1}^{m+2} - S_{n+1}^{m+1})}.$$

La probabilité q_n^m correspondante peut être calculée à partir de la condition de risque-neutralité (3.8), et les prix d'Arrow-Debreu sont calculés à partir des probabilités de transition et des prix d'Arrow-Debreu de l'étape précédente:

$$\lambda_{n+1}^i = \frac{1}{1+r\Delta t} (\lambda_n^{i-2} p_n^{i-2} + \lambda_n^{i-1} (1 - p_n^{i-1} - q_n^{i-1}) + \lambda_n^i q_n^i).$$

Au final, nous avons obtenu un algorithme de calcul *forward*:

Algorithme de calcul:
Initialiser: $\lambda_0^0 = 1$.
Pour $n = 0 \dots N-1$:
Calculer p_n^i et q_n^i, $i = 0, \dots, 2n$
Calculer λ_{n+1}^i, $i = 0, \dots, 2n+2$

Exemple 2. Supposons que la volatilité implicite observée est de la forme $I(K) = 0.15 \times \frac{100}{K}$ pour toutes maturités (la monnaie est à 100). Ceci nous permet de calculer les prix de toutes les options dont on a besoin. Les paramètres de l'arbre sont $T = 1$, $N = 2$ et $u = 1.2$. Les probabilités de transition (en petits caractères) et les prix d'Arrow-Debreu de l'arbre implicite sont montrés sur fig. 3.6.

Malgré l'attractivité théorique de la méthode, son implémentation met en lumière plusieurs difficultés. Premièrement, comme les prix d'options ne sont pas observés pour tous les noeuds de l'arbre, ils doivent être interpolés et le résultat final est très sensible à la méthode d'interpolation ainsi qu'au choix des noeuds de l'arbre. Ensuite, même si les données d'entrée ne contiennent pas d'opportunités d'arbitrage, les probabilités de transition peuvent être négatives (ou bien $p + q > 1$), ce qui conduit à des opportunités d'arbitrage dans les résultats du calcul. Il peut être assez difficile de trouver une structure de l'arbre qui donne des probabilités de transition positives. Pour résoudre ce deuxième problème, dans la section suivante on présentera une méthode permettant de construire un *processus de diffusion risque-neutre* dès que l'ensemble de prix observés ne présente pas d'opportunité d'arbitrage.

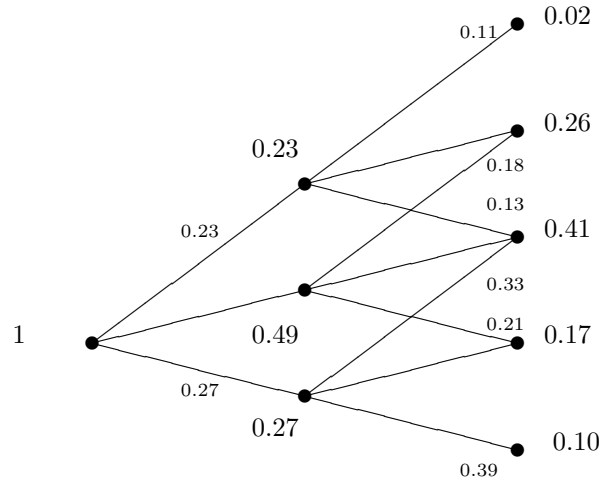


Figure 3.6: Arbre implicite avec les probabilités de transition (en petits caractères) et les prix Arrow-Debreu; voir exemple 2.

3.4 Diffusion implicite de Dupire

Rappelons que dans un modèle de volatilité locale le prix d'une option qui paie $h(S_T)$ à l'instant T satisfait

$$r_t C(t,S) = \frac{\partial C}{\partial t} + r_t S \frac{\partial C}{\partial S} + \frac{1}{2}\sigma(t,S)^2 S^2 \frac{\partial^2 C}{\partial S^2}, \quad C(T,S) = h(S). \tag{3.9}$$

Ceci est une équation *backward*, car on lui associe une condition terminale et l'équation se résout sur l'intervalle $[t,T]$ dans la direction $T \to t$. Cette équation est satisfaite par le prix de toute option européenne, pas nécessairement call ou put. Elle permet de calculer le prix d'une option en fonction de la date d'observation t et de la valeur actuelle du sous-jacent S_t.

Notre objectif maintenant est de trouver une fonction de volatilité locale $\sigma(t,S)$ qui reproduit, à une date donnée, les prix observés des calls pour tous les strikes et toutes les maturités. Equation (3.9) ne permet pas de reconstruire la volatilité locale en écrivant

$$\sigma^2(t,S) = \frac{rC - \frac{\partial C}{\partial t} - rS\frac{\partial C}{\partial S}}{\frac{1}{2}S^2\frac{\partial^2 C}{\partial S^2}},$$

car à une date donnée, les valeurs de t et S sont fixées, et on ne peut pas calculer les dérivées partielles. La solution à ce problème a été donnée par Bruno Dupire [27] qui a proposé une méthode pour calculer $\sigma(t,S)$ à partir d'une observation des prix d'options (pour tous strikes et maturités) à une date donnée. Il a démontré que dans un modèle à volatilité locale, les prix de calls $C(t,S_t,T,K)$ vérifient l'équation forward (avec condition initiale):

$$\frac{\partial C}{\partial T} = \frac{1}{2}\sigma^2(T,K)K^2\frac{\partial^2 C}{\partial K^2} - rK\frac{\partial C}{\partial K}, \quad C(t,S_t,t,K) = (S_t - K)^+$$

Cette équation s'applique uniquement aux prix des options Call, qui sont cette fois considérés comme fonctions du strike K et de la maturité T. Comme à une date donné on peut observer les prix d'options de plusieurs strikes et maturités, cette équation peut être utilisée pour recalculer la fonction de volatilité $\sigma(\cdot,\cdot)$ à partir des prix d'options. Ce résultat implique que dans un modèle de volatilité locale, la fonction de volatilité σ peut être retrouvée de façon unique avec:

$$\sigma(T,K) = \sqrt{2\frac{\frac{\partial C}{\partial T} + rK\frac{\partial C}{\partial K}}{K^2\frac{\partial^2 C}{\partial K^2}}} \tag{3.10}$$

Comme on l'a déjà observé dans le cas des arbres implicites, le fait qu'on puisse retrouver de façon unique un processus markovien continu à partir des prix d'options européennes n'implique pas qu'il n'y a pas d'autres

modèles (non markoviens ou non continus) qui évaluent les options européennes de la même façon. La connaissance des prix des options euro détermine les distributions marginales du processus, mais la loi du processus ne se limite pas à ces distributions marginales.

Théorème 1. *Soit $(S_t)_{t_0\leq t}$ une solution de l'EDS*

$$\frac{dS_t}{S_t} = rdt + \sigma(t,S_t)dW_t, \quad S_{t_0} = S_0.$$

Supposons que

1. *S_t est de carré intégrable:*
$$E\left[\int_{t_0}^{T} S_t^2 dt\right] < \infty, \quad \forall T$$
2. *Pour chaque $t > t_0$, la variable aléatoire S_t a une densité $p(t,x)$, continue sur $(t_0,\infty)\times(0,\infty)$.*
3. *Le coefficient de diffusion $\sigma(t,x)$ est continu sur $(t_0,\infty)\times(0,\infty)$.*

Alors la fonction de prix d'une option call

$$C(T,K) = e^{-r(T-t_0)}E[(S_T-K)^+].$$

satisfait l'équation de Dupire

$$\frac{\partial C}{\partial T} = \frac{1}{2}\sigma^2(T,K)K^2\frac{\partial^2 C}{\partial K^2} - rK\frac{\partial C}{\partial K}, \quad (T,K)\in[t_0,\infty)\times[0,\infty) \tag{3.11}$$

avec condition initiale $C(t_0,K) = (S_0-K)^+$.

Preuve. La démonstration repose sur l'application de la formule d'Itô à la semimartingale $e^{-rt}(S_t-K)^+$. Puisque la fonction $f(x) = x^+$ n'est pas C^2, la formule d'Itô classique ne s'applique pas directement. Une solution possible [28] consiste à utiliser la formule de Meyer-Itô pour les fonctions convexes [52]. Ici, nous adoptons une autre approche qui consiste à régulariser f en introduisant la fonction

$$f_\varepsilon(x) = \frac{(x+\varepsilon/2)^2}{2\varepsilon}1_{-\varepsilon/2\leq x\leq\varepsilon/2} + x1_{x>\varepsilon/2}.$$

Il est clair que f_ε est 2 fois différentiable et différente de f seulement si $|x| < \varepsilon/2$. De plus, on a

$$f'_\varepsilon(x) = \frac{x+\varepsilon/2}{\varepsilon}1_{-\varepsilon/2\leq x\leq\varepsilon/2} + 1_{x>\varepsilon/2}, \qquad f''_\varepsilon(x) = \frac{1}{\varepsilon}1_{-\varepsilon/2\leq x\leq\varepsilon/2}.$$

L'application de la formule d'Itô standard à $e^{-rt}f_\varepsilon(S_t-K)$ entre T et $T+\theta$ donne

$$\begin{aligned} e^{-r(T+\theta)}f_\varepsilon(S_{T+\theta}-K) - e^{-rT}f_\varepsilon(S_T-K) &= -r\int_T^{T+\theta} e^{-rt}f_\varepsilon(S_t-K)dt \\ &+ \int_T^{T+\theta} e^{-rt}f'_\varepsilon(S_t-K)dS_t + \frac{1}{2}\int_T^{T+\theta} e^{-rt}f''_\varepsilon(S_t-K)\sigma^2(t,S_t)S_t^2dt. \end{aligned} \tag{3.12}$$

En prenant l'espérance de chaque terme dans (3.12) sous l'hypothèse 1 ci-dessus, on trouve

$$\begin{aligned} e^{-r(T+\theta)}E[f_\varepsilon(S_{T+\theta}-K)] - e^{-rT}E[f_\varepsilon(S_T-K)] &= -r\int_T^{T+\theta} e^{-rt}E[f_\varepsilon(S_t-K)]dt \\ &+ \int_T^{T+\theta} e^{-rt}E[f'_\varepsilon(S_t-K)S_t]rdt \\ &+ \frac{1}{2}\int_T^{T+\theta} e^{-rt}\frac{1}{\varepsilon}\int_{-\varepsilon/2}^{\varepsilon/2}(K+x)^2\sigma^2(t,K+x)p(t,K+x)dt, \end{aligned} \tag{3.13}$$

En utilisant l'hypothèse 2 ci-dessus, on peut passer à la limite $\varepsilon \to 0$:

$$\begin{aligned}
&C(T+\theta,K) - C(T,K)\\
&= -r\int_T^{T+\theta} C(t,K)dt + r\int_T^{T+\theta} e^{-rt}E[S_t 1_{S_t\geq K}]dt\\
&\quad + \frac{1}{2}\int_T^{T+\theta} e^{-rt}\sigma^2(t,K)K^2p(t,K)dt\\
&= rK\int_T^{T+\theta} e^{-rt}P[S_t\geq K]dt + \frac{1}{2}\int_T^{T+\theta} e^{-rt}\sigma^2(t,K)K^2p(t,K)dt.
\end{aligned}$$

En divisant les deux parties par θ et en passant à la limite $\theta \to 0$, ceci donne

$$\frac{\partial C}{\partial T} = rKe^{-rT}P[S_T\geq K] + \frac{1}{2}e^{-rT}\sigma^2(T,K)K^2p(T,K).$$

Finalement, en observant que

$$e^{-rT}P[S_T\geq K] = -\frac{\partial C}{\partial K} \quad \text{et} \quad e^{-rT}p(T,K) = \frac{\partial^2 C}{\partial K^2},$$

on trouve l'équation de Dupire. □

Modèle à volatilité locale comme projection markovienne Soit $(S_t)_{t_0\leq t}$ un processus d'Itô général de la forme

$$\frac{dS_t}{S_t} = rdt + \sigma_t dW_t, \quad S_{t_0} = S_0$$

avec σ_t aléatoire, et supposons que $\sigma^2(t,x) = E[\sigma_t^2|S_t = x]$ vérifie les hypothèses du théorème ci-dessus. En suivant la preuve étape par étape, il est facile de se convaincre que les prix des options dans le modèle associé à (S_t) vérifient l'équation de Dupire avec volatilité $\sigma(t,x)$. Autrement dit, la fonction de volatilité $\sigma^2(t,x) = E[\sigma_t^2|S_t = x]$ définit une diffusion markovienne qui a les mêmes prix d'options européennes et donc les mêmes lois marginales que (S_t). Cette diffusion s'appelle la *projection markovienne* de (S_t).

Vision EDP de la formule de Dupire La formule de Dupire peut être vue comme une conséquence de l'équation Fokker-Planck pour la densité du processus. Soit $\mu_t(S)$ la densité de S_t. Alors, sous des hypothèses de régularité qu'on ne détaille pas ici,

$$\frac{\partial \mu_t}{\partial t} - \frac{1}{2}\frac{\partial^2}{\partial S^2}\left(\sigma^2 S^2\mu_t\right) + \frac{\partial}{\partial S}\left(rS\mu_t\right) = 0, \quad \mu_0(S) = \delta_{S_0}(S) \tag{3.14}$$

Cette équation s'appelle l'équation de Fokker-Planck ou l'équation forward de Kolmogorov pour le processus (S_t). Remarquons tout de même que les hypothèses nécessaires pour ce résultat sont bien plus fortes que celles du théorème 1: par exemple, ici on a besoin que la densité soit C^2.

Les prix des calls s'obtiennent en intégrant la densité risque-neutre deux fois:

$$Call(0,S_0,T,K) = e^{-rT}\int_K^\infty dS\int_S^\infty dS'\mu_T(S')$$

Pour retrouver l'équation de Dupire, on intègre donc l'équation de Fokker-Planck (3.14) deux fois terme par terme et multiplie par e^{-rT}:

- Premier terme:

$$e^{-rT}\int_K^\infty dS\int_S^\infty dS'\frac{\partial\mu_T(S')}{\partial T} = e^{-rT}\frac{\partial}{\partial T}(e^{rT}C(T,K)) = rC(T,K) + \frac{\partial C}{\partial T}.$$

- Deuxième terme:

$$-e^{-rT}\int_K^\infty dS\int_S^\infty dS'\frac{1}{2}\frac{\partial^2}{\partial S^2}(\sigma^2 S'^2\mu_t(S')) = -\frac{1}{2}\sigma^2(T,K)K^2\mu_T(K)e^{-rT}$$
$$= -\frac{1}{2}\sigma^2(T,K)K^2\frac{\partial^2 C}{\partial K^2}.$$

- Troisième terme:

$$e^{-rT}\int_K^\infty dS\int_S^\infty dS'\frac{\partial}{\partial S'}(rS'\mu_T(S')) = -e^{-rT}\int_K^\infty rS\mu_T(S)dS$$
$$= -\int_K^\infty rS\frac{\partial^2 C}{\partial K^2}dS = r\int_K^\infty \frac{\partial C}{\partial K}dS + rK\frac{\partial C}{\partial K} = rK\frac{\partial C}{\partial K} - rC(T,K).$$

En rassemblant les trois termes, on trouve (3.11).

Théorème 1 permet de retrouver le coefficient de volatilité à partir d'un ensemble complet de prix de calls à une date donnée, *si on sait que ces prix proviennent d'un modèle de volatilité locale.* Il ne permet pas directement de répondre à la question suivante: étant donné un système de prix de calls $(C(T,K))_{T\geq 0,K\geq 0}$, est-ce qu'il existe un modèle de diffusion continu permettant de reproduire ces prix? Pour appliquer la formule de Dupire (3.11) on a besoin au moins de supposer $\frac{\partial^2 C}{\partial K^2} > 0$ et $\frac{\partial C}{\partial T} + rK\frac{\partial C}{\partial K} \geq 0$. Ces contraintes correspondent aux contraintes d'arbitrage de positivité d'un *butterfly spread* et d'un *calendar spread* respectivement (voir section 1.4). Cependant, il se peut que même si la volatilité $\sigma(t,S)$ existe, elle ne conduit pas à un processus de Markov qui satisfait les trois hypothèses du théorème 1. Le résultat suivant donne une condition suffisante sous laquelle un processus de diffusion a une densité.

Proposition 2 (théorème 2.3.1 dans [51]). *Soit $(X_t, t\in[0,T])$ la solution d'une EDS*

$$dX_t = \mu(t,X_t)dt + \sigma(t,X_t)dW_t.$$

où les coefficients sont globalement Lipschitz et de croissance sous-linéaire, et soit

$$S = \inf\{t>0 : \int_0^t 1_{\sigma(s,X_s)>0}ds > 0\}\wedge T.$$

Alors, pour tout $0<t\leq T$, la loi de X_t conditionnellement à $\{t>S\}$ est absolument continu par rapport à la mesure de Lebesgue.

Par exemple, si $\sigma(0,X_0)>0$, le processus a une densité.

Exemple 3. Figure 3.7 montre les résultats d'application de la formule de Dupire aux données artificielles (gauche) et aux prix réels d'options sur l'indice S&P 500. Alors que sur les données simulées, la formule de Dupire permet de retrouver un surface de volatilité locale qui paraît cohérent, la performance pour les données réelles n'est pas satisfaisante pour plusieurs raisons:

- Les prix de marché ne sont pas connus pour tous les strikes et toutes les maturités. Ils doivent donc être interpolés et le résultat final sera très sensible à la méthode d'interpolation utilisée.
- Dû à la nécessité de calculer la deuxième dérivée de la fonction de prix d'option $C(T,K)$, les petites erreurs de données conduisent à des très grands erreurs sur la solution (problème mal posé). On reviendra sur ce point dans chapitre 6.

Lien entre volatilité locale et volatilité implicite La formule de Dupire (3.11) peut être réécrite en terme de volatilités implicites du marché, en observant que pour toute option on a

$$C(T,K) = C^{BS}(T,K,I(T,K)),$$

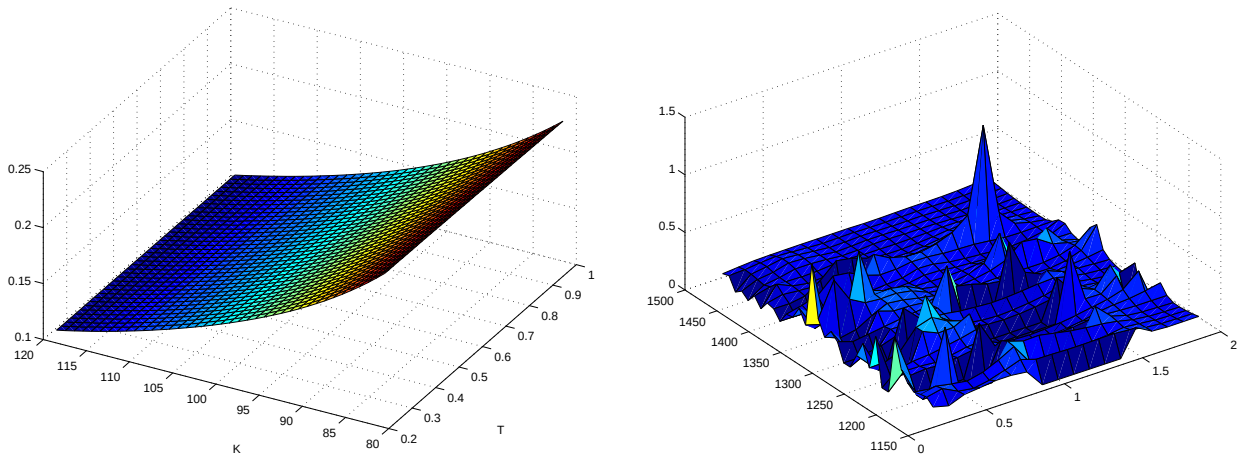


Figure 3.7: Exemple de diffusion implicite. Gauche: données artificielles: la volatilité implicite est de la forme $I(K) = 0.15 \times \frac{100}{K}$ pour toutes maturités ($S_0 = 100$). Droite: données d'options sur S&P 500, interpolation par splines.

où $C^{BS}(T, K, \sigma)$ dénote la formule Black-Scholes pour le prix d'une call de volatilité σ et $I(T, K)$ est la volatilité implicite observée pour maturité T et strike K. En substituant cette expression dans (3.11), on trouve

$$\begin{aligned}\sigma^2(T,K) &= 2\frac{\frac{\partial C^{BS}}{\partial T} + \frac{\partial C^{BS}}{\partial \sigma}\frac{\partial I}{\partial T} + rK\left(\frac{\partial C^{BS}}{\partial K} + \frac{\partial C^{BS}}{\partial \sigma}\frac{\partial I}{\partial K}\right)}{K^2\left(\frac{\partial^2 C^{BS}}{\partial K^2} + 2\frac{\partial^2 C^{BS}}{\partial K \partial \sigma}\frac{\partial I}{\partial K} + \frac{\partial^2 C^{BS}}{\partial \sigma^2}\left(\frac{\partial I}{\partial K}\right)^2 + \frac{\partial C^{BS}}{\partial \sigma}\frac{\partial^2 I}{\partial K^2}\right)} \\ &= \frac{\frac{I}{T} + 2\frac{\partial I}{\partial T} + 2rK\frac{\partial I}{\partial K}}{K^2\left(\frac{1}{K^2 I T} + 2\frac{d_1}{KI\sqrt{T}}\frac{\partial I}{\partial K} + \frac{d_1 d_2}{I}\left(\frac{\partial I}{\partial K}\right)^2 + \frac{\partial^2 I}{\partial K^2}\right)},\end{aligned} \tag{3.15}$$

avec la notation habituelle

$$d_{1,2} = \frac{\log\left(\frac{S}{Ke^{-rT}}\right) \pm \frac{1}{2}I^2 T}{I\sqrt{T}}.$$

Dans un premier temps supposons que la volatilité implicite ne dépend pas de strike (absence de smile). Dans ce cas la volatilité locale ne dépend pas non plus de strike et la formule (3.15) se réduit à

$$\sigma^2(T) = I^2(T) + 2I(T)T\frac{\partial I}{\partial T},$$

d'où

$$I^2(T) = \frac{\int_0^T \sigma^2(s)ds}{T},$$

la volatilité implicite est donc égale à la moyenne quadratique de la volatilité locale sur la durée de vie de l'option. Ce résultat peut être partiellement généralisé au cas où la volatilité dépend de strike [33].

Pour continuer l'étude de l'équation (3.15), on fera un changement de variable pour passer du strike K au *log-moneyness* $x = \log(S/K) + rT$, avec $I(T, K) = \tilde{I}(T, x)$. L'équation (3.15) devient alors

$$2\tilde{I}T\frac{\partial \tilde{I}}{\partial T} + \tilde{I}^2 - \sigma^2\left(1 - \frac{x}{\tilde{I}}\frac{\partial \tilde{I}}{\partial x}\right)^2 - \sigma^2 \tilde{I}T\frac{\partial^2 \tilde{I}}{\partial x^2} + \frac{1}{4}\sigma^2\tilde{I}^2T^2\left(\frac{\partial \tilde{I}}{\partial x}\right)^2 = 0.$$

En supposant que I et ses dérivées restent finis lorsque $T \to 0$, on a, dans cette limite,

$$\tilde{I}^2(0,x) = \sigma^2(0,x)\left(1 - \frac{x}{\tilde{I}}\frac{\partial \tilde{I}}{\partial x}\right)^2.$$

Cette équation différentielle se résout explicitement:

$$I(0,x) = \left\{ \int_0^1 \frac{dy}{\sigma(0,xy)} \right\}^{-1}. \tag{3.16}$$

Nous avons donc démontré que, *dans la limite de très courte maturité, la volatilité implicite est égale à la moyenne harmonique des volatilités locales,* un résultat établi par Berestycki et Busca [6]. Lorsque la volatilité locale $\sigma(0,x)$ est différentiable en $x = 0$, équation (3.16) permet de démontrer que (le calcul est laissé au lecteur)

$$\frac{\partial \tilde{I}(0,0)}{\partial x} = \frac{1}{2} \frac{\partial \sigma(0,0)}{\partial x},$$

la pente à la monnaie de la volatilité locale est égale, pour les courtes maturités, à 2 fois la pente à la monnaie de la volatilité implicite.

Résumé des principaux points critiques du modèle de volatilité locale et de l'approche de Dupire

- La surface de volatilité locale obtenue par la formule de Dupire est très sensible à la méthode d'interpolation des données
- Le seul facteur de risque est le sous-jacent; impossible de prendre en compte le risque de volatilité
- La surface de volatilité locale calibrée n'est pas stable au cours de temps
- La dynamique de volatilité préconisée par le modèle ne correspond pas à celle observée sur le marché
- Calibration ne peut pas être effectué de façon cohérente avec l'estimation historique (marché complet)

Exercices pour ce chapitre

Exercice 2 (Pas vraiment un exercice mais un problème difficile)**.** Donner un exemple de processus qui n'est pas une diffusion markovienne continue, et pour lequel la formule de Dupire permet de construire une diffusion markovienne continue qui donne les mêmes prix des options européennes.

Chapitre 4

Les modèles à volatilité stochastique

En observant les trajectoires de volatilité (fig. 4.1), on comprend immédiatement que ce processus n'est pas à 100% corrélé avec le sous-jacent et que donc cela a un sens de modéliser la volatilité comme un processus stochastique à part entière, sans se limiter à des fonctions déterministes de sous-jacent. Cette classe de modèles est connue sous le nom de modèles à volatilité stochastique. Ils ont comme avantage de non seulement expliquer le phénomène de smile (ce qu'on peut faire déjà dans les modèles à volatilité locale) mais aussi conduire à une dynamique réaliste pour le prix et la surface de volatilité implicite.

Une approche possible est de modéliser la volatilité σ_t par une diffusion continue (solution d'une EDS), mais on peut aussi envisager des processus de Markov discontinus, par exemple, un processus qui ne prend que deux valeurs σ_1 (volatilité basse) et σ_2 (volatilité haute). Le modèle de processus de diffusion, auquel on se limite ici, a un avantage de faciliter l'introduction d'une corrélation instantanée $-1 \leq \rho \leq 1$ entre le sous-jacent et la volatilité:

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t \tag{4.1}$$

$$d\sigma_t = a_t dt + b_t dW'_t, \quad d\langle W, W' \rangle_t = \rho dt \tag{4.2}$$

Les coefficients a_t et b_t sont choisis pour que la volatilité σ_t soit un processus stochastique positif. Souvent, on impose également la condition de stationnarité (retour à la moyenne). Figure 4.2 montre une trajectoire possible du sous-jacent et du processus de volatilité.

Le modèle (4.2) correspond en général à un marché incomplet: le risque de volatilité (risque de vega) ne peut pas être éliminé par une stratégie de trading en actions seulement. Cependant, on n'a ici que deux facteurs de risque (W et W'), ce qui fait que le risque de vega est complètement éliminé par une stratégie de couverture utilisant des actions et un autre actif risqué liquide (par exemple, une option).

4.1 Equations de pricing et de couverture

Dans cette section on se place dans un cadre markovien:

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t \tag{4.3}$$

$$d\sigma_t = a_t dt + b_t dW'_t, \qquad d\langle W, W' \rangle_t = \rho dt \tag{4.4}$$

$$\frac{dB_t}{B_t} = r dt,$$

où $a_t = a(t, \sigma_t, S_t)$, $b_t = b(t, \sigma_t, S_t)$. En présence de deux sources de risque on aura besoin de deux actifs risqués pour la couverture. On supposera donc qu'il existe un actif liquide coté au prix

$$C_t^0 = C^0(t, \sigma_t, S_t).$$

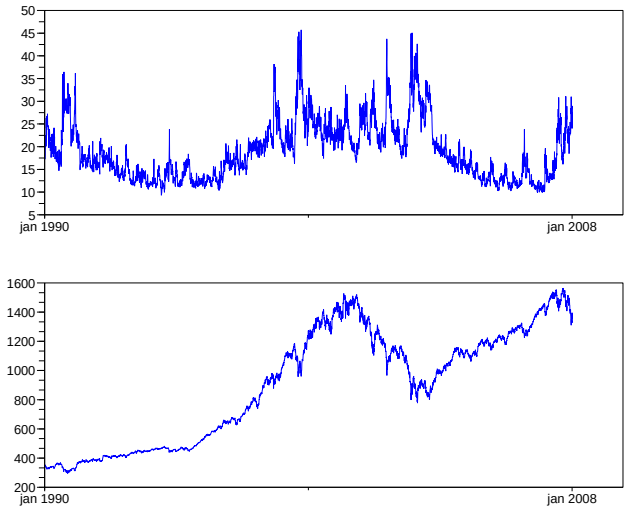


Figure 4.1: Haut: l'indice VIX (volatilité des options sur S&P 500). Bas: cours de l'indice S&P 500

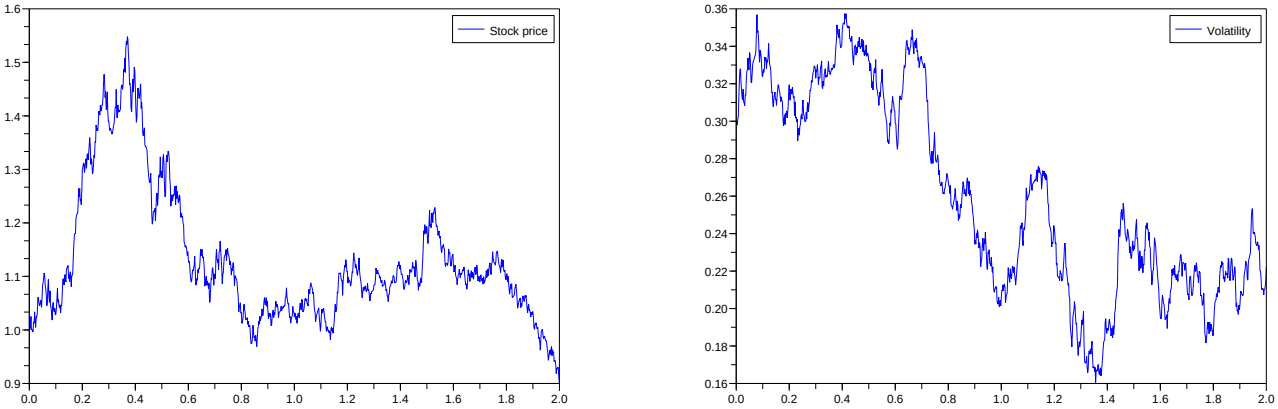


Figure 4.2: Prix de l'action (gauche) et trajectoires de volatilité dans un modèle à volatilité stochastique.

où la fonction déterministe $C^0(t,\sigma,S)$ est connue et satisfait $\frac{\partial C^0(t,\sigma,S)}{\partial\sigma} > 0$ pour tout t,σ,S (dans ce cas on dit que l'actif C^0 complète le marché). De plus on aura besoin de l'hypothèse $-1 < \rho < 1$ et de toutes les hypothèses techniques sur l'existence et régularité des solutions des EDS (4.3)–(4.4) que l'on ne donne pas ici, mais qui sont détaillées dans [53].

Soit V_t la valeur du portefeuille autofinançant contenant δ_t actions et ω_t unités de C^0. Par la condition d'autofinancement et la formule d'Itô,

$$\begin{aligned} dV_t &= (V_t - \delta_t S_t - \omega_t C_t^0) r dt + \delta_t dS_t + \omega_t dC_t^0 \\ &= (V_t - \delta_t S_t - \omega_t C_t^0) r dt \\ &\qquad + (\delta_t + \omega_t \frac{\partial C^0}{\partial S}) dS_t + \omega_t \frac{\partial C^0}{\partial \sigma} d\sigma_t + \omega_t \mathcal{L}_t C^0 dt \qquad (4.5) \\ \text{avec} \quad \mathcal{L}_t &= \frac{\partial}{\partial t} + \frac{1}{2} S_t^2 \sigma_t^2 \frac{\partial^2}{\partial S^2} + \frac{1}{2} b_t^2 \frac{\partial^2}{\partial \sigma^2} + S_t \sigma_t b_t \rho \frac{\partial^2}{\partial S \partial \sigma} \end{aligned}$$

Essayons de trouver une stratégie autofinançante qui réplique une fonction déterministe C (qui donnera le prix de l'option à couvrir par la suite): $V_t = C(t,\sigma_t,S_t)$. En appliquant la formule d'Itô, on trouve:

$$dV_t = \mathcal{L}_t C dt + \frac{\partial C}{\partial S} dS_t + \frac{\partial C}{\partial \sigma} d\sigma_t,$$

et pour l'égalité entre cette expression et (4.5), les ratios de couverture et la fonction C doivent satisfaire

$$\omega_t = \frac{\partial C/\partial\sigma}{\partial C^0/\partial\sigma} \qquad (4.6)$$

$$\delta_t = \frac{\partial C}{\partial S} - \omega_t \frac{\partial C^0}{\partial S} \qquad (4.7)$$

$$\mathcal{L}_t C - rC + rS_t \frac{\partial C}{\partial S} = \frac{\partial C}{\partial \sigma} \frac{\mathcal{L}_t C^0 - rC^0 + rS_t \frac{\partial C^0}{\partial S}}{\partial C^0/\partial\sigma} \qquad (4.8)$$

Le deuxième facteur à droite ne dépend pas de l'option qu'on cherche à couvrir, mais seulement de l'actif "universel" qu'on a choisi initialement pour la couverture. Soit

$$\lambda(t,\sigma,S) = -\frac{\mathcal{L}_t C^0 - rC^0 + rS\frac{\partial C^0}{\partial S}}{\partial C^0/\partial\sigma}$$

Nous avons démontré par réplication que le prix d'une option européenne de payoff $h(S_T)$ satisfait:

$$\begin{aligned} &\mathcal{L}_t C - rC + rS\frac{\partial C}{\partial S} + \lambda(t,\sigma,S)\frac{\partial C}{\partial\sigma} = 0, \quad C(T,\sigma,S) = h(S) \qquad (4.9) \\ \text{avec} \quad &\mathcal{L}_t = \frac{\partial}{\partial t} + \frac{1}{2}S^2\sigma^2\frac{\partial^2}{\partial S^2} + \frac{1}{2}b_t^2\frac{\partial^2}{\partial\sigma^2} + S\sigma b_t\rho\frac{\partial^2}{\partial S\partial\sigma} \end{aligned}$$

C'est la généralisation de l'équation de Black-Scholes aux modèles de volatilité stochastique. La couverture avec des actions et des actifs risqués d'un autre type avec des ratios (4.6)–(4.7) s'appelle la couverture en delta-vega (risque de vega = risque de volatilité). Graphique 4.3 illustre la performance de cette stratégie dans le modèle de Heston (voir ci-dessus) en comparant dynamiquement la valeur de l'option avec le portefeuille de couverture correspondant. On voit qu'effectivement, la couverture avec deux actifs élimine complètement le risque.

Evaluation risque-neutre L'EDP de pricing (4.9) est une équation en 3 dimensions qui n'est pas toujours facile à résoudre. Il serait plus satisfaisant d'exprimer le prix d'une option européenne dans un modèle de volatilité stochastique comme espérance, pour le calculer, par exemple, avec une méthode de Monte Carlo. Les prix des actifs cotés sont des espérances actualisés de leurs flux terminaux sous la probabilité risque-neutre, sous laquelle toute option satisfait

$$dC_t = rC_t dt + \text{martingale}.$$

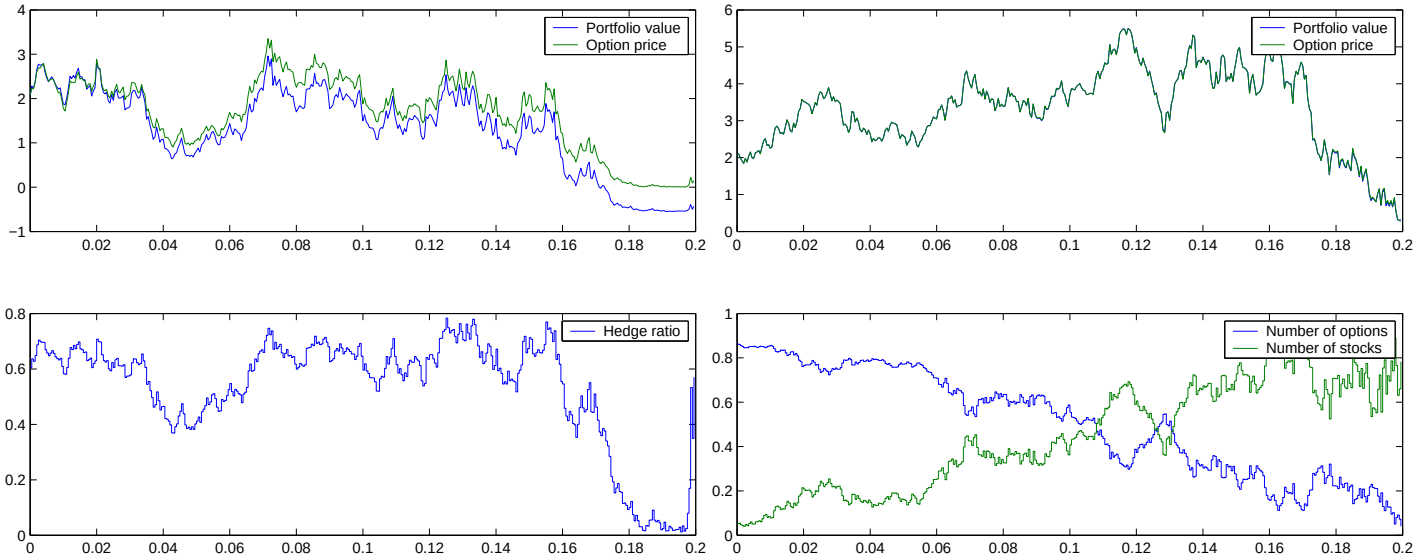


Figure 4.3: Couverture en vega. Gauche: couverture en delta seulement dans un marché à volatilité stochastique. Droite: couverture en delta-vega.

Or, en utilisant la formule d'Itô et (4.6)–(4.8), on trouve

$$dC_t = rC_t dt + \frac{\partial C}{\partial S}(dS_t - rS_t dt) + \frac{\partial C}{\partial \sigma}(d\sigma_t - \lambda_t dt),$$

donc, sous la probabilité risque-neutre Q on aura

$$\frac{dS_t}{S_t} = rdt + \sigma_t d\tilde{W}_t$$
$$d\sigma_t = \lambda_t dt + b_t d\tilde{W}'_t,$$

où $\tilde{W}$ et $\tilde{W}'$ sont Q-mouvements browniens avec $d\langle \tilde{W}, \tilde{W}' \rangle_t = \rho dt$. La "fonction universelle" λ, que nous avons rencontré pour la première fois dans l'équation de pricing (4.9) joue donc le rôle de la dérive de volatilité sous la probabilité risque-neutre.

Sous Q, le prix d'une option européenne de payoff h satisfait

$$C(t, \sigma, S) = E[e^{-r(T-t)} h(S_T) | \sigma_t = \sigma, S_t = S]$$

Sous la condition $\sigma_t > 0$ et $b_t > 0$, le modèle (4.3)-(4.4) peut être réécrit

$$\frac{dS_t}{S_t} = (r + \beta_t \sigma_t)dt + \sigma_t dW_t$$
$$d\sigma_t = (\lambda_t + \phi_t b_t)dt + b_t dW'_t,$$

Le coefficient β_t qui permet de passer de la probabilité risque-neutre à la probabilité historique est traditionnellement appelé la *prime de risque*. Par analogie, on appelle souvent ϕ_t la *prime de risque de volatilité*.

Exercice 3. Sous l'hypothèse que les primes de risque β_t et ϕ_t sont bornées, montrer comment on peut construire la probabilité risque-neutre Q à l'aide du théorème de Girsanov.

Remarque 1. Une différence fondamentale entre β_t et ϕ_t est que β_t est facile à estimer comme $\frac{\mu_t - r}{\sigma_t}$ alors que pour estimer ϕ_t on devrait estimer λ_t à partir de données historiques, ce qui n'est en général pas possible. Une solution consiste à spécifier une forme paramétrique pour λ_t qui sera calibrée directement sous la probabilité risque-neutre, aux prix d'options cotés sur le marché. On reviendra sur ce point dans la discussion des paramétrisations de volatilité stochastique.

4.2 Comportement asymptotique de volatilité implicite

Cas de volatilité de volatilité petite Dans cette section, supposons pour simplifier que la volatilité est indépendante du mouvement brownien W qui dirige le sous-jacent. On notera

$$V_T = \int_0^T \sigma_t^2 dt$$

la variance intégrée. Grâce à l'indépendance, en insérant l'espérance conditionnelle par rapport à la trajectoire de la volatilité, on trouve que le prix d'une option européenne vérifie

$$\text{Call}(T,K) = E[\text{Call}_{BS}(T,K,\sqrt{V_T/T})] = E[\overline{\text{Call}}_{BS}(T,K,V_T)],$$

où on a posé $\overline{\text{Call}}_{BS}(T,K,V) := \text{Call}_{BS}(T,K,\sqrt{V/T})$, $\forall V > 0$. Supposons maintenant que V_T est proche de $E[V_T]$, c'est-à-dire, $\text{Var}V_T \ll E[V_T]^2$. Par une décomposition en série de Taylor de la fonction $\overline{\text{Call}}_{BS}$, autour du point $E[V_T]$, on trouve

$$\text{Call}(T,K) \approx \overline{\text{Call}}_{BS}(T,K,E[V_T]) + \frac{1}{2}\text{Var}V_T \frac{\partial^2}{\partial V^2}\overline{\text{Call}}_{BS}(T,K,E[V_T]), \tag{4.10}$$

où

$$\frac{\partial^2}{\partial V^2}\overline{\text{Call}}_{BS}(T,K,V) = \frac{Sn(d_1)}{4V^{3/2}}\left\{\frac{m^2}{V} - \frac{V}{4} - 1\right\}, \quad m = \log\frac{S}{Ke^{-rT}}.$$

Par ailleurs, lorsque $\delta\sigma \ll \sigma$,

$$\text{Call}_{BS}(T,K,\sigma+\delta\sigma) \approx \text{Call}_{BS}(T,K,\sigma) + \delta\sigma\frac{\partial \text{Call}_{BS}(T,K,\sigma)}{\partial\sigma}.$$

En comparant ces deux développements, on trouve que la volatilité implicite $I(T,K)$ vérifie

$$\begin{aligned} I(T,K) &\approx \sqrt{E[V_T/T]} + \frac{1}{2}\text{Var}V_T \frac{\frac{\partial^2}{\partial V^2}\overline{\text{Call}}_{BS}(T,K,E[V_T])}{\frac{\partial}{\partial\sigma}\text{Call}_{BS}(T,K,\sqrt{E[V_T/T]})}, \\ &= \sqrt{E[V_T/T]} + \frac{1}{8T^{1/2}E[V_T]^{3/2}}\text{Var}V_T\left\{\frac{m^2}{E[V_T]} - \frac{E[V_T]}{4} - 1\right\}. \end{aligned} \tag{4.11}$$

Cas d'une option proche de l'echéance Pour analyser le comportement de la volatilité implicite pour les options de courte maturité en utilisant les résultats du paragraphe précedent, il faut connaitre les propriétés de $E[V_T]$ et $\text{Var}V_T$ lorsque $T \to 0$. Supposons que le processus de la variance instantanée $v_t = \sigma_t^2$ suit l'EDS homogène

$$dv_t = \alpha(v_t)dt + \beta(v_t)dW_t',$$

où le mouvement brownien W' est toujours supposé indépendant de W. Les fonctions α et β sont régulières et vérifient les hypothèses nécessaires pour l'existence d'une solution forte. Soit

$$u(t,v) = E\left[\int_t^T v_s ds | v_t = v\right].$$

Le processus

$$M_t = E\left[\int_0^T v_s ds \middle| \mathcal{F}_t\right] = \int_0^t v_s ds + u(t,v_t)$$

est alors une martingale, ce qui implique (formule d'Itô) que u vérifie l'équation

$$\frac{\partial u}{\partial t} + \alpha(v)\frac{\partial u}{\partial v} + \frac{1}{2}\beta^2(v)\frac{\partial^2 u}{\partial v^2} + v = 0, \quad u(T,v) \equiv 0. \tag{4.12}$$

On cherche à obtenir un développement de u en puissances de $T-t$. Puisque la condition terminale et les coefficients sont régulières (et la condition terminale est nulle),

$$u(t,v) = (T-t)u_1(v) + \frac{(T-t)^2}{2}u_2(v) + O((T-t)^3).$$

En substituant cette expression dans l'EDP (4.12), par identification des termes avec la même puissance de $(T-t)$, on trouve $u_1(v) = v$ et $u_2(v) = \alpha(v)$, d'où finalement

$$E[V_T] = u(0,v_0) = v_0 T + \frac{\alpha(v_0)T^2}{2} + O(T^3).$$

Le calcul de $\mathrm{Var}V_T$ suit la même logique. Soit

$$\bar{u}(t,v) = E\left[\left(\int_t^T v_s ds\right)^2 \middle| v_t = v\right].$$

Le processus

$$\bar{M}_t = E\left[\left(\int_0^T v_s ds\right)^2 \middle| \mathcal{F}_t\right] = \bar{u}(t,v_t) + \left(\int_0^t v_s ds\right)^2 + 2u(t,v_t)\int_0^t v_s ds$$

est alors une martingale. En appliquant la formule d'Itô, on trouve

$$\begin{aligned} d\bar{M}_t &= \left(\frac{\partial \bar{u}}{\partial t} + \alpha(v)\frac{\partial \bar{u}}{\partial v} + \frac{1}{2}\beta^2(v)\frac{\partial^2 \bar{u}}{\partial v^2}\right)dt + \frac{\partial \bar{u}}{\partial v}\beta(v)dW_t' \\ &\quad + 2\left[\left(v_t + \frac{\partial u}{\partial t} + \alpha(v)\frac{\partial u}{\partial v} + \frac{1}{2}\beta^2(v)\frac{\partial^2 u}{\partial v^2}\right)dt + \frac{\partial u}{\partial v}\beta(v)dW_t'\right]\int_0^t v_s ds + 2vu(t,v)dt \\ &= \left(\frac{\partial \bar{u}}{\partial t} + \alpha(v)\frac{\partial \bar{u}}{\partial v} + \frac{1}{2}\beta^2(v)\frac{\partial^2 \bar{u}}{\partial v^2} + 2vu(t,v)\right)dt + \left(\frac{\partial \bar{u}}{\partial v}\beta(v) + 2\frac{\partial u}{\partial v}\beta(v)\int_0^t v_s ds\right)dW_t', \end{aligned}$$

ce qui implique que $\bar{u}$ doit vérifier l'EDP

$$\frac{\partial \bar{u}}{\partial t} + \alpha(v)\frac{\partial \bar{u}}{\partial v} + \frac{1}{2}\beta^2(v)\frac{\partial^2 \bar{u}}{\partial v^2} + 2vu(t,v) = 0, \quad \bar{u}(T,v) \equiv 0. \tag{4.13}$$

Comme avant, posons

$$\bar{u}(t,v) = (T-t)\bar{u}_1(v) + \frac{(T-t)^2}{2}\bar{u}_2(v) + \frac{(T-t)^3}{3}\bar{u}_3(v) + O(T^4).$$

La substitution de cette forme dans (4.13) donne

$$\bar{u}_1(v) = 0, \quad \bar{u}_2(v) = 2v^2 \quad \text{et} \quad \bar{u}_3(v) = 3v\alpha(v) + \beta^2(v),$$

d'où

$$E[V_T^2] = v_0^2 T^2 + (v_0\alpha(v_0) + \beta^2(v_0)/3)T^3 + O(T^4) \quad \text{et} \quad \mathrm{Var}V_T = \frac{1}{3}\beta^2(v_0)T^3 + O(T^4).$$

L'hypothèse $\mathrm{Var}V_T \ll E[V_T]^2$, qui assure la validité de l'approximation (4.10) est donc respectée lorsque T est petit. Si de plus, l'option est proche de la monnaie ($|m| \ll 1$), on peut substituer les expressions pour $E[V_T]$ et $\mathrm{Var}V_T$ dans (4.11), ce qui donne

$$I(T,K) \approx \sqrt{v_0} + \frac{\beta^2(v_0)}{24v_0^{5/2}}m^2 + \left\{\frac{\alpha(v_0)}{4\sqrt{v_0}} - \frac{\beta^2(v_0)}{24v_0^{3/2}} - \frac{\beta^2(v_0)\alpha(v_0)}{48v_0^{7/2}}m^2\right\}T.$$

En particulier, pour $m = 0$, la volatilité implicite à la monnaie forward vérifie

$$I_{ATMF}(T) \approx \sqrt{v_0} + \left\{\frac{\alpha(v_0)}{4\sqrt{v_0}} - \frac{\beta^2(v_0)}{24v_0^{3/2}}\right\}T,$$

ce qui implique que la volatilité instantanée est la limite quand $T \to 0$ de la volatilité implicite à la monnaie forward. Pour $T = 0$ et m non nul (mais petit),

$$I(T = 0, K) \approx \sqrt{v_0} + \frac{\beta^2(v_0)}{24 v_0^{5/2}} m^2,$$

la volatilité implicite en $T = 0$ a localement une forme parabolique en m, centrée en $m = 0$. Ceci permet de calculer la *convexité* du smile aux courtes maturités:

$$\lim_{T\to 0} \frac{\partial^2 I}{\partial m^2}|_{m=0} = \frac{\beta^2(v_0)}{12 v_0^{5/2}}.$$

Des résultats de ce type ont des applications en calibration, puisqu'ils relient directement des quantités observables aux paramètres du modèle.

La forme localement symétrique de la volatilité implicite est une conséquence de l'hypothèse de l'indépendance des mouvements browniens du sous-jacent et volatilité ($\rho = 0$). Dans le cas $\rho \neq 0$, des résultats similaires ont été obtenus par Medvedev et Scaillet [50]. Ces auteurs modélisent directement la volatilité σ_t et non la variance instantanée:

$$d\sigma_t = a(\sigma_t)dt + b(\sigma_t)dW'_t, \quad d\langle W, W'\rangle_t = \rho dt.$$

Ceci correspond à prendre $\alpha(v) = b^2(\sqrt{v}) + 2a(\sqrt{v})\sqrt{v}$ et $\beta(v) = 2b(\sqrt{v})\sqrt{v}$ dans notre formulation. L'hypothèse que m doit être proche de zéro est interpretée en introduisant le moneyness renormalisé $z = \frac{m}{\sigma_0\sqrt{T}}$ et supposant que z reste fini losque $T \to 0$. Avec les méthodes de développement en série pour les EDP, similaires à celles qu'on vient de voir dans un cadre plus simple, Medvedev et Scaillet montrent alors que

$$I(T, z) = \sigma_0 + I_1(\sigma_0, z)\sqrt{T} + I_2(\sigma_0, z)T + O(T\sqrt{T})$$

avec

$$I_1(\sigma, z) = \frac{\rho b(\sigma) z}{2}$$

et

$$I_2(\sigma, z) = \frac{1}{6}\left\{\frac{b^2(\sigma)(1-\rho^2)}{\sigma} + \rho^2 b(\sigma)b'(\sigma)\right\} z^2 + \frac{a(\sigma)}{2} + \frac{\rho\sigma b(\sigma)}{4} + \frac{\rho^2 b^2(\sigma)}{24\sigma} + \frac{b^2(\sigma)}{12\sigma} - \frac{\rho^2 b(\sigma)b'(\sigma)}{6}.$$

Cette formule permet de calculer le skew et la convexité à la monnaie pour courtes maturités dans le cas $\rho \neq 0$:

$$\lim_{T\to 0} \frac{\partial I}{\partial m}|_{m=0} = \frac{\rho b(\sigma_0)}{2\sigma_0} = \frac{\rho\beta(v_0)}{4v_0},$$
$$\lim_{T\to 0} \frac{\partial^2 I}{\partial m^2}|_{m=0} = \frac{1}{3\sigma_0^2}\left\{\frac{b^2(\sigma_0)(1-\rho^2)}{\sigma_0} + \rho^2 b(\sigma_0)b'(\sigma_0)\right\} = \frac{\beta^2(v_0)(1-2\rho^2)}{12 v_0^{5/2}} + \frac{\rho^2\beta(v_0)\beta'(v_0)}{6 v_0^{3/2}}.$$

La formule de Roger Lee pour les strikes extrêmes Dans cette section, qui n'est pas limitée aux modèles à volatilité stochastique, on montre que la condition d'absence d'arbitrage impose des contraintes sur la croissance de la volatilité implicite pour les très grands et les très petits strikes. Ces résultats ont été établis par Roger Lee [45] et généralisés dans [31].

Soit $k = \log \frac{Ke^{-rT}}{S}$ le log strike. On fixe l'échéance T et denote par $I(k)$ la volatilité implicite d'une option d'échéance T et log-strike k, et par $c^{BS}(k, \sigma)$ le prix Black-Scholes de cette option, si la volatilité est égale à σ. Dans tout modèle sans arbitrage, on a alors,

$$\limsup_{k\to+\infty} \frac{I^2(k)T}{k} \leq 2. \tag{4.14}$$

Pour démontrer cette inégalité il suffit de voir qu'il existe $k^* < \infty$ tel que pour tout $k > k^*$,

$$c^{BS}(k, I(k)) < c^{BS}\left(k, \sqrt{\frac{2k}{T}}\right).$$

Or,

$$\lim_{k\to\infty} c^{BS}(k, I(k)) = 0$$

par convergence dominée, alors que

$$\lim_{k\to\infty} c^{BS}\left(k, \sqrt{\frac{2k}{T}}\right) = \lim_{k\to\infty} S\left\{N(0) - e^k N(-\sqrt{2k})\right\} = \frac{S}{2}$$

par le règle de l'Hopital.

Le résultat plus général de Roger Lee établit le lien entre la valeur de $\limsup_{k\to+\infty} \frac{I^2(k)T}{k}$ et le nombre des moments finis du processus de prix S_T. Dans un premier temps on va étudier le comportement du prix d'un call. Soit $E[S_T^{p+1}] < \infty$ pour $p > 0$. Il est facile de voir que pour tout $S > 0$,

$$S - K \le \frac{S^{p+1}}{p+1}\left(\frac{p}{p+1}\right)^p \frac{1}{K^p} \quad \text{et donc} \quad (S-K)^+ \le \frac{S^{p+1}}{p+1}\left(\frac{p}{p+1}\right)^p \frac{1}{K^p}$$

En prenant l'espérance actualisée de chaque coté,

$$C(K) \le e^{-rT}\frac{E[S^{p+1}]}{p+1}\left(\frac{p}{p+1}\right)^p \frac{1}{K^p}$$

et donc

$$C(K) = O(K^{-p}), \quad K \to \infty \tag{4.15}$$

Avec (4.14) et (4.15), on a finalement

Théorème 2 (Roger Lee)**.**

(i) Soit

$$\tilde{p} = \sup\{p : E[S_T^{1+p}] < \infty\} \quad et \quad \beta_R = \limsup_{k\to+\infty} \frac{I^2(k)T}{k}.$$

Alors $\beta_R \in [0,2]$ *et*

$$\tilde{p} = \frac{1}{2\beta_R} + \frac{\beta_R}{8} - \frac{1}{2}.$$

(ii) Soit

$$\tilde{q} = \sup\{q : E[S_T^{-q}] < \infty\} \quad et \quad \beta_L = \limsup_{k\to-\infty} \frac{I^2(k)T}{|k|}.$$

Alors $\beta_L \in [0,2]$ *et*

$$\tilde{q} = \frac{1}{2\beta_L} + \frac{\beta_L}{8} - \frac{1}{2}.$$

Preuve. Nous donnerons ici seulement une idée de la preuve de la première partie du théorème, qui permet de comprendre comment apparait le lien entre le nombre de moments finis et le comportement de la volatilité implicite. Le lecteur peut consulter [45] pour la preuve complète. Soit $f_\pm(\beta) := \frac{1}{\beta} + \frac{\beta}{4} \pm 1$. Alors,

$$c^{BS}(k, \sqrt{\beta k/T}) = S(N(-\sqrt{f_-(\beta)k}) - e^k N(-\sqrt{f_+(\beta)k})). \tag{4.16}$$

En remarquant que

$$\lim_{x\to-\infty} \frac{N(-x)}{|x|^{-1}n(x)} = 1,$$

on montre facilement que

$$c^{BS}(k, \sqrt{\beta k/T}) \sim C\frac{e^{-f_-(\beta)k/2}}{\sqrt{k}} \tag{4.17}$$

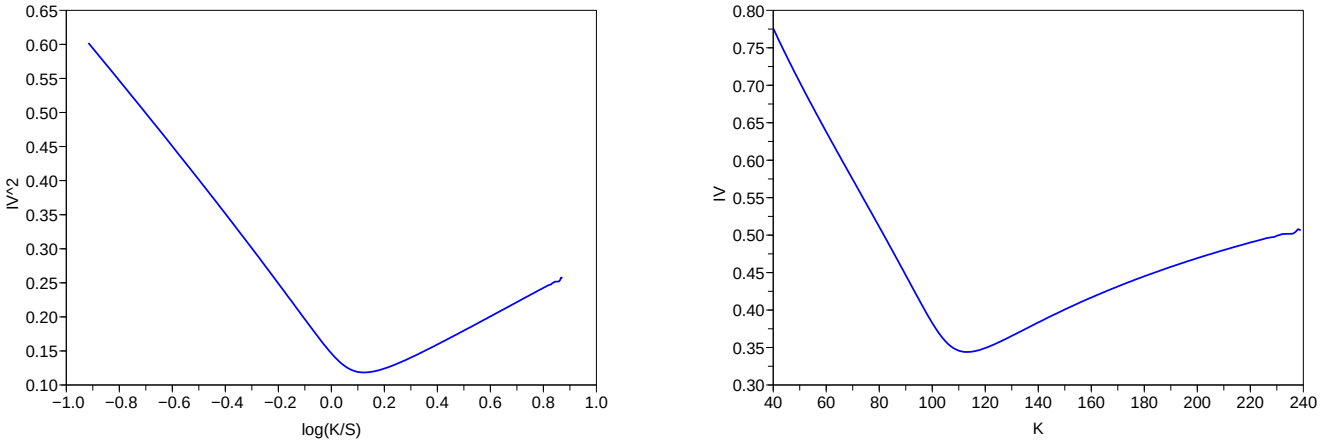


Figure 4.4: Comportement asymptotique du smile dans le modèle de Heston pour les valeurs extrêmes du strike. Les paramètres sont: $\rho = -0.5$, $v_0 = \theta = 0.04$, $\delta = 0.5$ et $k = 2$, temps jusqu'à maturité: $T = 50$ jours.

lorsque $k \to \infty$, où C est une constante. Soit $p < \tilde{p}$. On a alors en utilisant (4.15):

$$\frac{c^{BS}(k, I(k))}{c^{BS}(k, \sqrt{\beta k/T})} = O(e^{-(p - f_-(\beta)/2)k}\sqrt{k}).$$

Ceci montre que $\tilde{p} > \frac{f_-(\beta)}{2}$ implique $\beta_R \leq \beta$, et puisque la fonction $f_-(\beta)$ est decroissante, que $\tilde{p} \leq f_-(\beta_R)/2$. □

Le résultat de Roger Lee montre que le carré de la volatilité, en fonction du log-strike, est asymptotiquement linéaire. Ce type de comportement est illustré sur le graphique 4.4.

4.3 Estimation de volatilité

Dans la section précédente nous avons vu qu'en théorie, dans un modèle à volatilité stochastique avec 2 facteurs de risque, toute option est répliqué par un portefeuille contenant des actions et un actif risqué additionnel. En réalité, pour calculer les ratios de couverture (4.6) et (4.7) et le prix de l'option, il faut connaître la volatilité instantanée σ_t qui n'est pas directement observable. On peut approcher σ_t par la volatilité moyenne sur une période de longueur T, qui est estimable à partir de données historiques, mais

- La variance de l'estimateur de volatilité moyenne décroît avec T
- Le biais de l'approximation de σ_t par la volatilité moyenne croit avec T

En général, même avec des données haute fréquence, on ne peut connaître σ_t qu'avec une précision de 10% au mieux, ce qui introduit un risque supplémentaire (risque de modèle) dans la procedure de couverture. Dans cette section on se propose de regarder les méthodes d'estimation de volatilité moyenne à partir de données historiques.

Estimateurs de variance réalisée Soit (S_t) un processus de prix avec volatilité stochastique:

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t.$$

Alors

$$d \log S_t = \left(\mu_t - \frac{\sigma_t^2}{2}\right) dt + \sigma_t dW_t \tag{4.18}$$

et la variation quadratique de $\log S_t$ satisfait

$$\langle \log S \rangle_t = \int_0^t \sigma_s^2 ds.$$

D'un autré coté, la variation quadratique d'une sémimartingale est définie par

$$\langle X \rangle_t = \lim_{N\to\infty} \sum_{i=1}^{N} (X_{\frac{i}{N}t} - X_{\frac{i-1}{N}t})^2,$$

où la convergence a lieu presque sûrement. Avec la notation

$$r(t,h) \equiv \log \frac{S_{t+h}}{S_t}$$

pour le log-rendement du processus de prix, on a donc

$$\int_0^t \sigma_s^2 ds = \lim_{N\to\infty} \sum_{i=1}^{N} r\left(\frac{i-1}{N}t, \frac{t}{N}\right)^2.$$

Ceci suggère qu'on peut estimer la variance intégrée $\int_0^t \sigma_s^2 ds$ par la *variance réalisée*

$$VR_t^N \equiv \sum_{i=1}^{N} r\left(\frac{i-1}{N}t, \frac{t}{N}\right)^2.$$

Pour analyser le comportement de cet estimateur, on fera deux hypothèses simplificatrices:

- La volatilité (σ_t) est indépendante du mouvement brownien qui dirige le sous-jacent (W_t).
- La dérive μ_t est déterministe.

Ces hypothèses ne sont pas nécessaires, pour voir comment on peut s'en affranchir, cf. [2].

On écrira E^σ pour l'espérance conditionnelle par rapport à la trajectoire de volatilité. Sous cette espérance conditionnelle,

$$R_i \equiv r\left(\frac{i-1}{N}t, \frac{t}{N}\right)$$

est une variable aléatoire Gaussienne de variance et espérance

$$v_i = \int_{\frac{i-1}{N}t}^{\frac{i}{N}t} \sigma_s^2 ds \quad \text{et} \quad m_i = \int_{\frac{i-1}{N}t}^{\frac{i}{N}t} (\mu_s - \sigma_s^2/2) ds$$

respectivement, et les variables R_i et R_j sont indépendantes pour tout $i \neq j$.

Dans un premier temps, on analyse le biais de l'estimateur de variance réalisée. Un calcul direct montre

$$\begin{aligned} E^\sigma[VR_t^N] &= \sum_{i=1}^{N} E^\sigma[R_i^2] = \sum_{i=1}^{N} (v_i + m_i^2) \\ &\approx \int_0^t \sigma_s^2 ds + \frac{t}{N} \int_0^t (\mu_s - \sigma_s^2/2)^2 ds, \end{aligned}$$

le biais de l'estimateur est donc de l'ordre de $O(t/N)$. Pour la variance conditionnelle de l'estimateur on a

$$\begin{aligned} \mathrm{Var}^\sigma[VR_t^N] &= \sum_{i=1}^{N} E^\sigma[(R_i^2 - v_i - m_i^2)^2] = \sum_{i=1}^{N} (2v_i^2 + 4v_i m_i^2) \\ &\approx \frac{2t}{N} \int_0^t \sigma_s^4 ds + \frac{4t^2}{N^2} \int_0^t \sigma_s^2 (\mu_s - \sigma_s^2/2)^2 ds. \end{aligned}$$

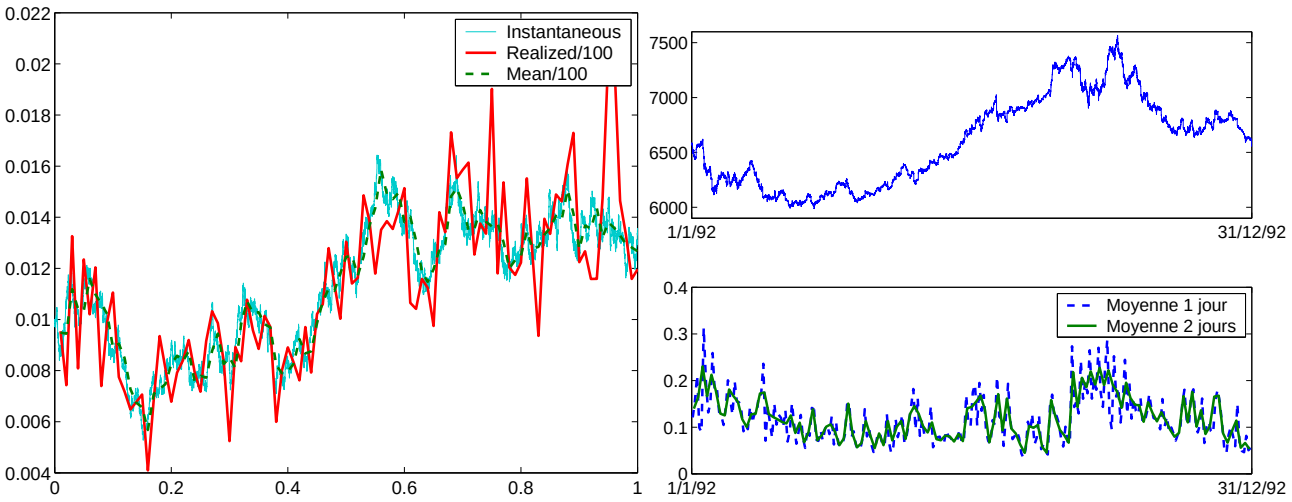


Figure 4.5: Estimateur de variance réalisée. Graphique de gauche: données simulées. Graphique de droite: données réelles. Haut: taux de change DM/USD. Bas: volatilité moyenne journalière, estimée sur les données de fréquence 5 min.

Le terme dominant dans l'erreur de l'estimateur est donc dû à la variance et l'écart type de l'erreur est, au premier ordre, égal à

$$\sqrt{\frac{2t}{N}\int_0^t \sigma_s^4 ds}.$$

Ceci s'explique par le fait que l'estimateur de variance réalisée est un estimateur complètement non-paramétrique, il ne nécessite aucune hypothèse sur le modèle et n'introduit donc presque pas de biais mais au prix d'une variance assez élevée.

Exemple 4. Soit $t = 1$ jour, $N = 100$ (une donnée toutes les 5 minutes). Alors pour un niveau moyen de volatilité de 10%, on a

$$\frac{2t}{N}\int_0^t \sigma_s^4 ds \sim 5 \times 10^{-6},$$

ce qui correspond à une erreur relative de 15%. D'un autre coté, l'écart type typique de la volatilité instantanée sur 1 jour est de 5×10^{-3} (paramètres de l'article de Heston [40]). La précision relative de l'estimation de la volatilité instantanée par l'estimateur de variance réalisée est donc, dans ce cas de 16%.

L'estimation de volatilité par variance réalisée est illustrée par fig. 4.5. A gauche, sur les données simulées, on compare l'estimateur de variance réalisée sur 1 jour avec la volatilité instantanée et la volatilité moyenne sur 1 jour, c'est-à-dire $\sqrt{\frac{1}{t}\int_0^t \sigma_s^2 ds}$. On voit qu'ici, l'écart entre la volatilité instantanée et la volatilité moyenne est petit, mais l'estimateur de variance réalisée introduit beaucoup de bruit. Sur le graphique de droite, on applique le même estimateur avec des fenêtres d'un jour et de deux jours aux données réelles. Ici également, l'importante différence entre les estimateurs pour les deux fenêtres montre que l'estimateur de variance réalisée introduit beaucoup de bruit dans le résultat final.

Estimation de volatilité avec un modèle ARCH auxiliaire Cette méthode d'estimation de volatilité introduit un lissage supplémentaire, ce qui permet de diminuer la variance de l'estimateur, au prix d'un biais plus grand. On se place dans le cadre d'un modèle *discret* à volatilité stochastique:

$$R_i = \mu_i + \sigma_i \varepsilon_i. \tag{4.19}$$

Ici, $\{\varepsilon_i\}$ est une suite de variables aléatoires adaptée à une filtration discrète $\{\mathcal{F}_t\}$ et telle que pour tout i, ε_{i+1} est indépendant de $\mathcal{F}_i$ et de loi $N(0,1)$. Les suites $\{\mu_i\}$ et $\{\sigma_i\}$ sont supposées $\{\mathcal{F}_t\}$-prévisibles (ceci revient à

dire que σ_i et μ_i sont $\mathcal{F}_{i-1}$-mesurables). Dans ce cas, le rendement continu $r((i-1)h, h)$ a la même loi pour tout i que le rendement discret $R_i = \mu_i + \sigma_i \varepsilon_i$.

Remarque 2. Sous l'hypothèse de dérive déterministe et volatilité indépendante du sous-jacent, le modèle continu (4.18) observé avec fréquence h peut être réécrit sous cette forme, en prenant une suite $\{\varepsilon_i\}$ de variables indépendantes de loi $N(0,1)$, et en posant

$$\mathcal{F}_t := \sigma\{\varepsilon_1, \dots, \varepsilon_t\} \vee \sigma\{\sigma_s : s \le (t+1)h\}.$$
$$\mu_i = \int_{(i-1)h}^{ih} (\mu_s - \sigma_s^2/2) ds, \quad \sigma_i = \sqrt{\int_{(i-1)h}^{ih} \sigma_s^2 ds}.$$

Dans le modèle (4.19),

$$\sigma_i^2 = E[(R_i - \mu_i)^2 | \mathcal{F}_{i-1}].$$

Ceci conduit à l'idée d'approcher σ_i^2 par une moyenne mobile des carrées des observations précédentes

$$\sigma_i^2 \approx \omega + \alpha_1 (R_{i-1} - \mu_{i-1})^2 + \dots + \alpha_n (R_{i-n} - \mu_{i-n})^2. \tag{4.20}$$

Cette méthode peut donc être considéré comme une généralisation de l'estimateur de variance réalisée. D'un autre coté, il existe une classe de modèles économétriques de volatilité, où elle a exactement la forme (4.20). Il s'agit de modèles de type ARCH (autoregressive conditional heteroscedasticity) [30]. Dans un modèle ARCH(n), on suppose que le processus de rendements discrets est décrit par (4.19), avec une volatilité donnée par

$$\sigma_i^2 = \omega + \alpha_1 (R_{i-1} - \mu_{i-1})^2 + \dots + \alpha_n (R_{i-n} - \mu_{i-n})^2. \tag{4.21}$$

Les moyennes μ_i peuvent être constantes ou intégrer une éventuelle saisonnalité.

Dans ce modèle tout l'aléa provient donc de la suite $\{\varepsilon_i\}$, et il n'y a pas de facteurs de risque cachés ce qui facilite l'estimation. Pour estimer un modèle ARCH(n), on peut dans un premier temps utiliser la méthode de maximum de vraisemblance pour déduire les paramètres ω, α_1,…,α_n puis filtrer la volatilité avec l'équation (4.21). Par exemple, dans le modèle ARCH(1) avec $\mu_i \equiv \mu$ la densité jointe de $(R_1, \dots, R_T)$ satisfait

$$p(R_1, \dots, R_T | \theta) = p(R_T | R_{T-1}, \theta) \dots p(R_2 | R_1, \theta) p(R_1),$$

où $\theta = (\mu, \omega, \alpha_1)$ est le vecteur des paramètres inconnus et

$$\begin{aligned} p(R_i | R_{i-1}, \theta) &= \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(R_i - \mu)^2}{2\sigma_i^2} \right) \\ &= \frac{1}{\sqrt{2\pi(\omega + \alpha_1 (R_{i-1} - \mu)^2)}} \exp\left(-\frac{(R_i - \mu)^2}{2(\omega + \alpha_1 (R_{i-1} - \mu)^2} \right). \end{aligned}$$

La densité $p(R_1)$ peut être Gaussienne avec une valeur de volatilité fixée à priori. Les paramètres (μ, ω, α_1) peuvent donc être estimés avec

$$\hat{\theta} = \arg\max_\theta \log p(R_1, \dots, R_T | \theta) = \arg\max_\theta \left(-\sum_{i=1}^T \log \sigma_i - \sum_{i=1}^T \frac{(R_i - \mu)^2}{2\sigma_i^2} \right).$$

Pour construire un estimateur fiable de volatilité avec un modèle ARCH auxiliaire, on est obligé de choisir n suffisamment grand et donc d'estimer beaucoup de paramètres. Pour une description plus parcimonieuse, les modèles GARCH (generalized ARCH) ont été proposés [7], qui permettent à la volatilité σ_i de dépendre non seulement des rendements passés mais aussi des valeurs passées de volatilité elle-même. Le modèle GARCH(m,n) a donc la forme

$$\sigma_t^2 = \omega + \sum_{i=1}^n \alpha_i (R_{t-i} - \mu_{t-i})^2 + \sum_{j=1}^m \beta_j \sigma_{t-j}^2.$$

L'avantage de cette approche est que pour obtenir des bons résultats on a rarement besoin d'aller au-delà de GARCH(1,1).

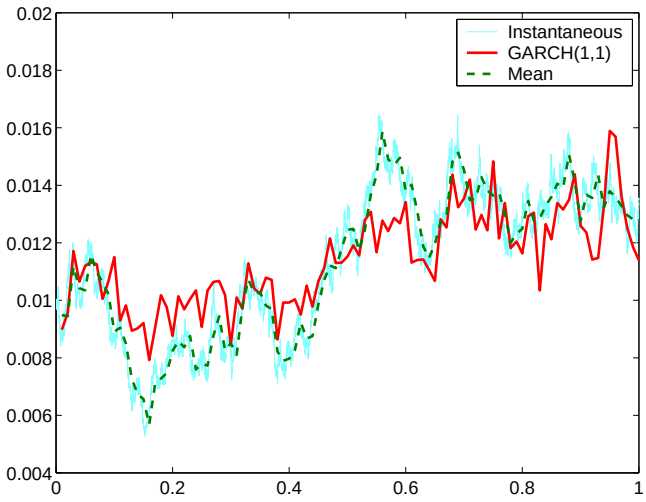


Figure 4.6: Estimation de volatilité avec un modèle GARCH(1,1) auxiliaire.

Pour estimer la volatilité dans un modèle à volatilité stochastique, une méthode possible consiste donc à substituer la volatilité stochastique par une volatilité de type ARCH ou GARCH. En le faisant, on, introduit un biais dû au fait qu'on n'utilise pas le bon modèle, mais en vue de l'équation (4.20) on peut espérer avoir une approximation à la vraie volatilité. Graphique 4.6 montre la performance de l'estimateur fondé sur un modèle GARCH(1,1) auxiliaire sur les mêmes données qui ont été utilisées pour tester l'estimateur de variance réalisée (fig. 4.5). On voit que maintenant, l'estimateur contient beaucoup moins de bruit, mais que la volatilité estimée est parfois différente de la vraie volatilité en raison d'un lissage indésirable par le modèle GARCH.

Estimation de la volatilité instantanée avec volatilité implicite Des résultats de la section précedente il ressort que la volatilité implicité à la monnaie forward (et la volatilité implicite à la monnaie) converge vers la volatilité instantanée $\sqrt{v_0}$ pour les petites maturités:

$$I_{ATM}(T) = \sqrt{v_0} + O(\sqrt{T}) \quad \text{(cas général)}$$
$$I_{ATM}(T) = \sqrt{v_0} + O(T) \quad (\rho = 0).$$

Dans les marchés liquides, on peut donc estimer la volatilité implicite simplement en interpolant la volatilité implicite à la monnaie. Les formules de la section précedente permettent d'analyser l'erreur d'une telle approximation.

4.4 Swaps de variance

Les swaps de variance sont des actifs contingents qui permettent de prendre des positions sur la volatilité (variance) de sous-jacent. Un swap de variance a un pay-off à l'échéance égal à

$$H_T = \frac{NA}{n} \sum_{i=1}^{n} \left(\log \frac{S_i}{S_{i-1}} \right)^2 - N\sigma_K^2,$$

où

- $A = 250$ est le nombre moyen de jours ouvrés dans l'année;
- N est le nominal du contral;
- n represente le nombre de jours ouvrés jusqu'a l'échéance T;
- σ_K est la volatilité 'strike'.

En d'autres mots, un swap de variance permet d'echanger un montant fixe $N\sigma_K^2$ contre un flux aléatoire égal à la variance réalisée du sous-jacent.

Dans un modèle à trajectoires continues, un swap de variance peut être répliqué par un portefeuille statique contenant des options européennes et un portefeuille dynamique contenant le sous-jacent [11]. Pour simplifier le traitement, on va approcher la somme dans le pay-off de produit par une intégrale:

$$H_T = N\left\{\frac{1}{T}\int_0^T \sigma_s^2 ds - \sigma_K^2\right\}.$$

Supposons que le sous-jacent S est décrit par

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t,$$

où μ_t et σ_t sont deux processus stochastiques. La formule d'Itô donne alors

$$\frac{1}{2}\int_0^T \sigma_s^2 ds = \int_0^T \frac{dS_t}{S_t} - \log\frac{S_T}{S_0}. \tag{4.22}$$

Pour répliquer la variance intégrée, il suffit alors de répliquer les deux termes dans la partie droite.

Soit V_t la valeur du portefeuille qui réplique $\int_0^T \frac{dS_t}{S_t}$ et soit Q une probabilité risque-neutre. On a alors

$$\begin{aligned} V_t &= e^{-r(T-t)} E\left[\int_0^T \frac{dS_t}{S_t}|\mathcal{F}_t\right] \\ &= e^{-r(T-t)}\left\{\int_0^t \frac{dS_s}{S_s} + r(T-t)\right\} \end{aligned}$$

d'où on deduit

$$dV_t = r(V_t - \delta_t S_t)dt + \delta_t dS_t \quad \text{avec} \quad \delta_t = \frac{e^{-r(T-t)}}{S_t}.$$

Le premier terme de la partie droite de (4.22) peut donc être répliqué par un portefeuille autofinançant consistant à investir $e^{-r(T-t)}$ en actions et ayant la valeur initiale $V_0 = rTe^{-rT}$.

Le deuxième terme dans la partie droite de (4.22) (le log-contrat) peut être répliqué par un portefeuille statique de calls et puts. Soit f une fonction C^2. Alors il est facile de demontrer que pour tous F, x positifs,

$$f(x) = f(F) + f'(F)(x-F) + \int_0^F f''(K)(K-x)^+ dK + \int_F^\infty f''(K)(x-K)^+ dK.$$

Pour calculer le prix d'une option européenne d'échéance T et de pay-off $f(S_T)$, on pose $F = S_0 e^{rT}$ et calcule l'espérance de $e^{-rT}f(S_T)$ sous la probabilité risque-neutre:

$$\text{Prix} = e^{-rT}f(F) + \int_0^F f''(K)P(T,K)dK + \int_F^\infty f''(K)C(T,K)dK,$$

où $P(T,K)$ est le prix d'un put et $C(T,K)$ le prix d'un call d'échéance T et strike K. En particulier, pour le log-contrat $f(x) = \log\frac{x}{S_0}$, on trouve

$$\text{Prix} = e^{-rT}rT - \int_0^F \frac{P(T,K)}{K^2}dK - \int_F^\infty \frac{C(T,K)}{K^2}dK.$$

En rajoutant ceci au prix du portefeuille de replication pour le premier terme de (4.22), on trouve que le prix du portefeuille de replication pour $\frac{1}{2}\int_0^T \sigma_s^2 ds$ est égal à

$$\int_0^F \frac{P(T,K)}{K^2}dK + \int_F^\infty \frac{C(T,K)}{K^2}dK.$$

Finalement, le σ_K qui annule la valeur du swap de variance satisfait

$$e^{-rT}\sigma_K^2 = \frac{2}{T}\int_0^F \frac{P(T,K)}{K^2}dK + \frac{2}{T}\int_F^\infty \frac{C(T,K)}{K^2}dK.$$

L'indice VIX VIX est l'indice de volatilité d'options sur S&P 500, publié par CBOE. Entre 1993 et 2003, cet indice était calculé comme la moyenne des volatilités implicites de 8 options les plus liquides, mais en 2003 la méthodologie a été changé, et le nouvel indice est calculé en prenant en compte toutes les options par la formule (4.4). Le rapport publié sur le site de CBOE (faites un recherche google sur "vix white paper") donne la nouvelle formule: on calcule

$$VIX_T^2 = \frac{2e^{rT}}{T}\sum_i \frac{\Delta K_i}{K_i^2} Q_i(K_i) - \frac{1}{T}\left[\frac{F}{K_0} - 1\right]^2,$$

pour les deux échéances les plus courtes T_1 et T_2, où

- Q_i est le prix de l'option (call ou put) qui est hors de la monnaie pour le strike K_i;
- F est le forward calculé par la parité call-put;
- K_0 est le plus grand strike qui est plus petit que le forward;
- ΔK_i est l'intervalle entre les strikes.

Le troisième terme est un terme de correction:

$$\begin{aligned}
&\frac{2}{T}\int_0^F \frac{P(T,K)}{K^2}dK + \frac{2}{T}\int_F^\infty \frac{C(T,K)}{K^2} \\
&= \frac{2}{T}\int_0^{K_0} \frac{P(T,K)}{K^2}dK + \frac{2}{T}\int_{K_0}^\infty \frac{C(T,K)}{K^2}dK + \frac{2}{T}\int_{K_0}^F \frac{P(K)-C(K)}{K^2}dK \\
&\approx \frac{2}{T}\int_0^{K_0} \frac{P(T,K)}{K^2}dK + \frac{2}{T}\int_{K_0}^\infty \frac{C(T,K)}{K^2}dK - \frac{e^{-rT}}{T}\frac{(F-K_0)^2}{K_0^2}.
\end{aligned}$$

Pour calculer l'indice VIX, on fait une interpolation linéaire entre VIX_{T_1} et VIX_{T_2} pour arriver à une maturité de 30 jours.

4.5 Paramétrisation de Heston

Dans sa version risque-neutre, le modèle de Heston [40] a la forme:

$$\begin{aligned}
\frac{dS_t}{S_t} &= rdt + \sigma_t dW_t \\
d\sigma_t^2 &= k(\theta - \sigma_t^2)dt + \delta\sigma_t dW_t', \quad d\langle W, W'\rangle_t = \rho dt \\
\text{où encore} \quad dv_t &= k(\theta - v_t)dt + \delta\sqrt{v_t}dW_t' \quad \text{avec} \quad v_t = \sigma_t^2
\end{aligned}$$

Le processus de la variance v_t est connu sous le nom de processus racine carrée. Il a été introduit par Cox, Ingersoll et Ross [19] pour modéliser le taux d'intérêt court. Ce processus reste positif et retourne à la moyenne θ avec vitesse caractéristique k. Le paramètre δ joue le rôle de volatilité de volatilité. Le comportement dynamique du processus dépend des grandeurs relatives des paramètres: si $\frac{2k\theta}{\delta^2} > 1$, le processus ne touche jamais zéro; dans le cas contraire il peut toucher zéro et réfléchir. Dans cette dernière situation, on rencontre des difficultés de simulation de ce processus: puisque la fonction $f(x) = \sqrt{x}$ n'est pas Lipschitz en zéro, le schéma d'Euler

$$v_{t+\Delta t} = v(t) + k(\theta - v_t)\Delta t + \delta\sqrt{v_t}\Delta W_t'$$

converge très lentement. De plus, même si le processus continu reste toujours positif, le processus discrétisé peut devenir négatif. Cette difficulté est généralement évitée en remplaçant $\sqrt{v_t}$ par $\sqrt{|v_t|}$ ou $\sqrt{v_t^+}$, mais ceci ne résout pas le problème de convergence. Pour voir comment améliorer la convergence avec un schéma d'ordre élevé, voir [33, 35]. Il existe également des algorithmes de simulation exacte (sans discrétisation) — voir [9] — mais ils ne sont pas toujours les plus rapides.

Pour mieux comprendre les propriétés du modèle, on peut réécrire le processus racine carrée comme

$$v_t = \theta + e^{-kt}(v_0 - \theta) + \delta \int_0^t e^{-k(t-s)} \sqrt{v_s} dW'_s.$$

Dans cette forme on voit immédiatement que le processus "oublie" sa valeur initiale et retourne vers la moyenne θ exponentiellement vite avec vitesse k. La moyenne du processus est

$$E[v_t] = e^{-kt}(v_0 - \theta) + \theta$$

et la variance

$$\mathrm{Var}[v_t] = \frac{\delta^2 \theta}{2k} + \frac{\delta^2 (v_0 - \theta)}{k} e^{-kt} + \frac{\delta^2 (\theta - 2v_0)}{2k} e^{-2kt}.$$

Si t est petit,

$$\mathrm{Var}[v_t] = v_0 \delta^2 t,$$

ce qui explique le terme "volatilité de volatilité".

Remarquons au passage que

$$\frac{1}{T} E\left[\int_0^T v_t dt\right] = \theta + (v_0 - \theta)\frac{1 - e^{-kT}}{kT},$$

ce qui permet de calculer facilement le prix d'un swap de variance.

Les paramètres dans l'article original de Heston [40] sont $k = 2$, $\theta = 0.01$, $\delta = 0.1$ et $v_0 = 0.01$. Pour ces valeurs, l'écart type de variance sur un jour ($T = 4 \cdot 10^{-3}$) est de $\approx 6 \cdot 10^{-4}$. Si on remplace la variance instantanée par la variance intégrée sur 1 jour, on fait donc une erreur relative de 6%.

Exercice 4. Calculer la variance de la variance moyenne $\mathrm{Var} I_T$ dans le modèle de Heston. Montrer que $\mathrm{Var} I_T \sim O(T)$ quand $T \to 0$.

Valorisation d'options européennes Section 5.8 décrit une méthode rapide pour calculer les prix des options européennes dans un modèle où on connaît explicitement la fonction caractéristique du log-prix. Le modèle de Heston appartient à la classe de modèles dits affines [25, 26], et pour tous les modèles dans cette classe, la fonction caractéristique est connue. Ici on explique la méthode de calcul pour le modèle de Heston seulement.

Le log-prix X_t est défini par $S_t = S_0 e^{rt + X_t}$ où le taux d'intérêt a été introduit pour simplifier les calculs par la suite. Le processus (X_t) satisfait l'EDS

$$dX_t = -\frac{1}{2}\sigma_t^2 dt + \sigma_t dW_t$$

Par le lemme d'Itô, la fonction caractéristique conditionnelle de X_T,

$$f(t, v, x) = E[e^{iuX_T} | X_t = x, v_t = v],$$

satisfait:

$$\frac{v}{2}\frac{\partial^2 f}{\partial x^2} + \rho\delta v \frac{\partial^2 f}{\partial x \partial v} + \frac{\delta^2 v}{2}\frac{\partial^2 f}{\partial v^2} - \frac{v}{2}\frac{\partial f}{\partial x} + k(\theta - v)\frac{\partial f}{\partial v} + \frac{\partial f}{\partial t} = 0,$$

avec condition finale $f(T, v, x) = e^{iux}$.

Cherchons la solution sous la forme :

$$f(t, v, x) = \Phi(t, v) e^{iux}. \tag{4.23}$$

Alors l'équation devient:

$$\frac{\delta^2 v}{2}\frac{\partial^2 \Phi}{\partial v^2} + (iu\rho\delta v + k(\theta - v))\frac{\partial \Phi}{\partial v} - \frac{v}{2}(u^2 + iu)\Phi + \frac{\partial \Phi}{\partial t} = 0,$$

avec condition finale $\Phi(T,v)=1$. Grâce à la structure affine des coefficients, on devine encore une fois la forme de la solution:

$$\Phi(t,v)=\exp\{\theta C(T-t)+vD(T-t)\}. \tag{4.24}$$

L'équation se décompose alors en 2 équations différentielles ordinaires:

$$D'(t)=\frac{\delta^2}{2}D^2+(iu\rho\delta-k)D-\frac{u^2+iu}{2},\quad D(0)=0 \tag{4.25}$$

$$C'(t)=kD,\quad C(0)=0 \tag{4.26}$$

Ces EDO pour C et D ont des solutions explicites:

$$D(t)=-\frac{u^2+iu}{\gamma\coth\frac{\gamma t}{2}+k-i\rho\delta u}, \tag{4.27}$$

$$C(t)=\frac{kt(k-i\rho\delta u-\gamma)}{\delta^2}-\frac{2k}{\delta^2}\ln\left(\frac{1}{2}(1+e^{-\gamma t})+\frac{k-i\rho\delta u}{2\gamma}(1-e^{-\gamma t})\right), \tag{4.28}$$

où $\gamma=\sqrt{\delta^2(u^2+iu)+(k-i\rho\delta u)^2}$. Finalement, en rassemblant équations (4.23)–(4.24), on trouve la fonction caractéristique non-conditionnelle de (X_t):

$$E[e^{iuX_T}]=e^{\theta C(T)+v_0D(T)}.$$

Remarque 3. L'écriture (4.23)–(4.24) des facteurs C et D de la fonction caractéristique de Heston est "optimisée" pour les calculs numériques: l'expression sous le signe de logarithme n'explose pas pour u grand, et reste dans le domaine de continuité de la fonction log (la fonction de logarithme complexe dans la plupart des bibliothèques numériques a une discontinuité sur le demi-axe $(-\infty,0]$). Voir [46] pour une discussion approfondie de ce problème

Explosions des moments dans le modèle de Heston L'analyse détaillée des équations (4.25) et (4.26) [42] montre que $E[e^{uX_T}]$ (et donc $E[S_T^u]$) est fini si et seulement si $T<T^*(u)$, où

$$T^*(u)=\begin{cases}+\infty, & \Delta(u)\geq 0\\ \dfrac{2}{\sqrt{-\Delta}}\left(\arctan\dfrac{\sqrt{-\Delta(u)}}{\rho\delta u-k}+\pi 1_{\rho\delta u-k<0}\right), & \Delta(u)<0,\end{cases}$$

$$\text{where}\quad \Delta(u)=(k-u\rho\delta)^2-\delta^2(u^2-u).$$

Cela a comme conséquence que les prix de certaines options dont le pay-off n'est pas à croissance sous-linéaire (de telles options existent en particulier dans le domaine de taux d'intérêt) peuvent exploser dans le modèle de Heston. Cette formule donne par ailleurs le moyen de calculer les asymptotiques de la volatilité implicite pour les valeurs extrêmes de strike dans le modèle de Heston (voir théorème 2), mais les formules qu'on obtient ne sont pas très explicites.

Volatilité implicite dans le modèle de Heston Les quatre graphiques de fig. 4.7 illustrent le rôle de différents paramètres du modèle de Heston pour la description du smile de volatilité implicite. On voit que les trois paramètres v_0, ρ et δ déterminent respectivement le niveau, la pente et la convexité du smile à une maturité donnée, alors que θ et k jouent sur la structure par terme de volatilité implicite. Graphique 4.8 montre une surface typique de volatilité implicite dans le modèle de Heston.

En utilisant les résultats de Medvedev et Scailler (voir section 4.2), on peut calculer les asymptotiques de skew et de la convexité de la volatilité implicite dans le modèle de Heston:

$$\lim_{T\to 0}\frac{\partial I(T,m)}{\partial m}|_{m=0}=\frac{\rho\delta}{2\sigma_0};\quad \lim_{T\to 0}\frac{\partial^2 I(T,m)}{\partial m^2}|_{m=0}=\frac{\delta^2(1-\rho^2)}{3\sigma_0^3};$$

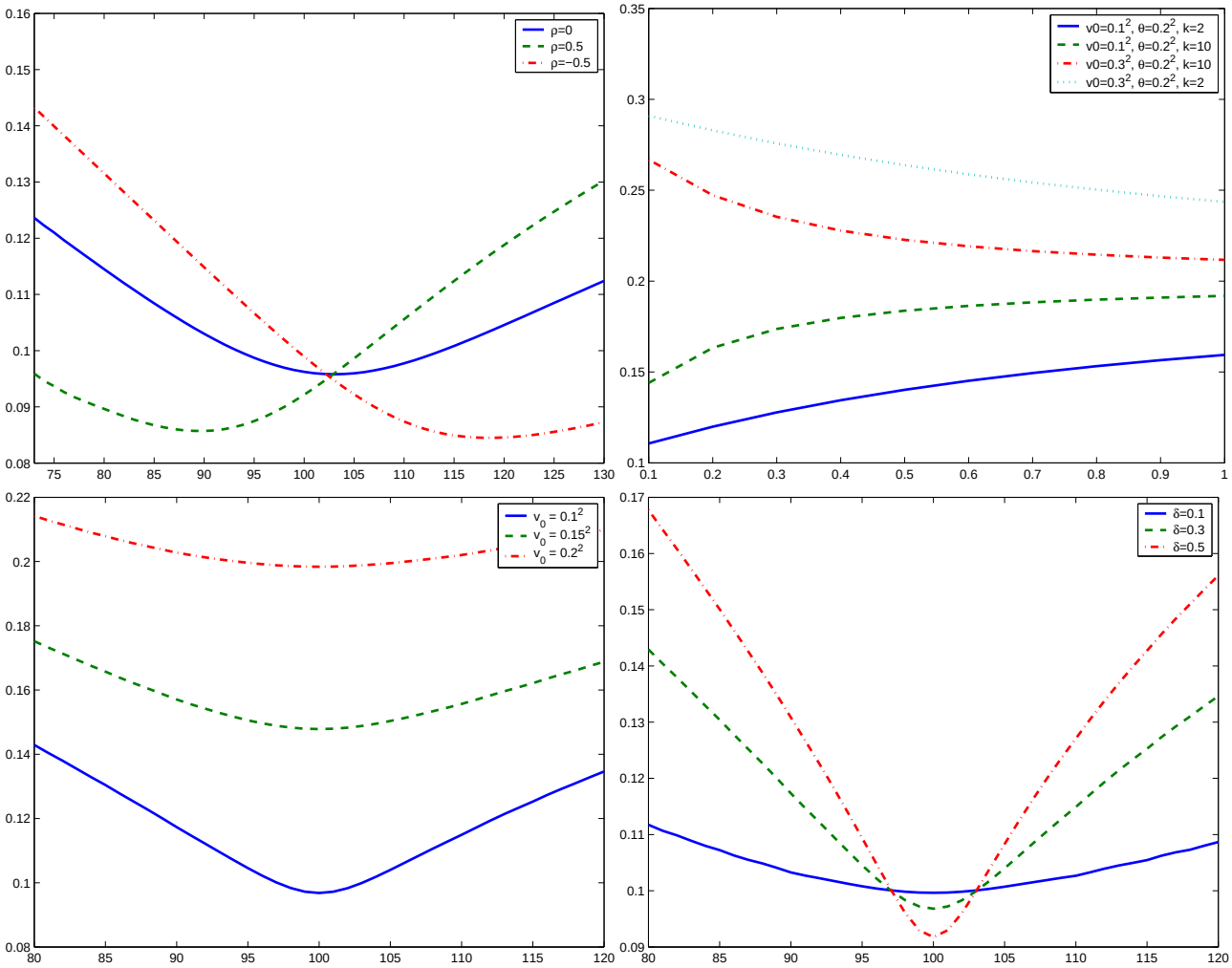


Figure 4.7: Influence de différents paramètres du modèle de Heston sur le comportement de la volatilité implicite. En haut à gauche: influence de la corrélation. En haut à droite: influence de k et θ sur la structure par terme. En bas à gauche: influence du niveau initial de volatilité v_0. En bas à droite: influence la volatilité de volatilité δ.

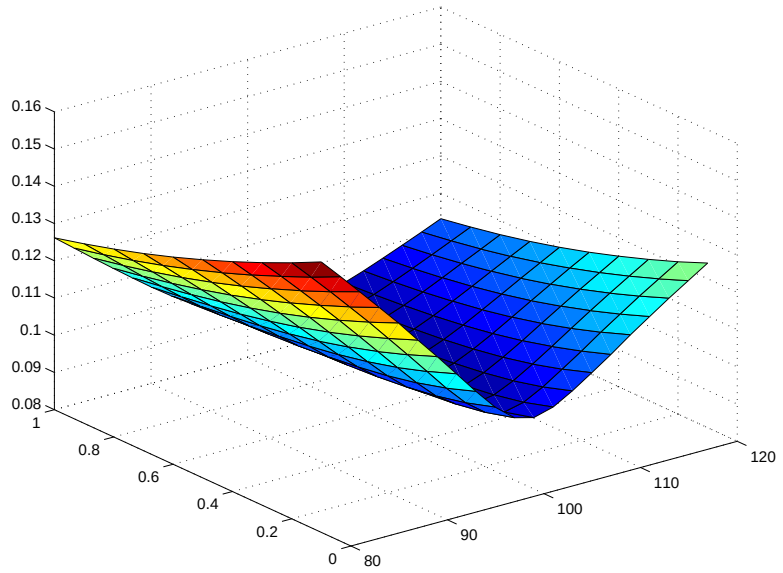


Figure 4.8: Surface de volatilité implicite dans le modèle de Heston.

Calibration du modèle de Heston Dans cette dernière section on présente un étude de cas de calibration du modèle de Heston aux données réelles. Nous avons pris les prix de clôture des calls et puts sur l'indice S&P 500 (source: Datastream). Le taux de dividende a été estimé avec la parité call-put à la monnaie. Seulement les options hors de la monnaie avec prix $> \$0.25$ utilisées et la calibration a été effectuée par moindres carrées avec points de départ différents pour éviter de tomber dans un minimum local. Les résultats sont représentés sur les trois graphiques du fig. 4.9. On voit que le modèle reproduit correctement la pente de la volatilité implicite aux grandes maturités (même pour plusieurs maturités en même temps), mais n'arrive pas à expliquer la convexité aux très courtes maturités. Ceci est du au fait que sur une échelle de temps courte la volatilité ne peut pas changer de manière significative, et le modèle se comporte comme le modèle Black-Scholes.

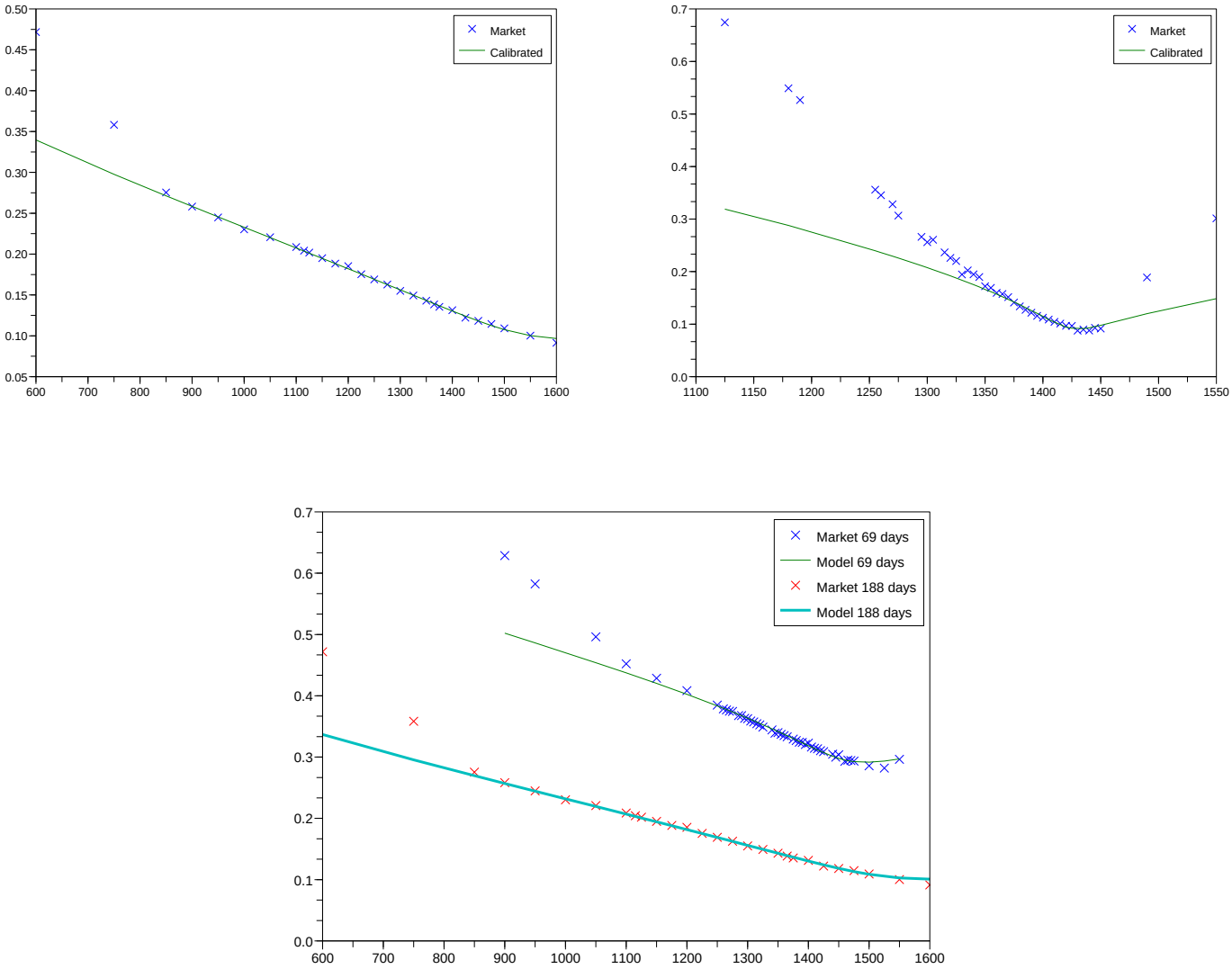


Figure 4.9: Calibration du modèle de Heston aux données de marché. En haut à gauche: Maturité 188 jours: le modèle reproduit correctement le skew du marché aux longues maturités. En haut à droite: Maturité 8 jours: le smile aux courtes maturités n'est pas calibré correctement. En bas: Calibration simultanée à 2 maturités: la structure par terme est modélisée correctement.

Chapitre 5

Les modèles avec sauts

Dans ce chapitre, on considère les processus stochastiques avec sauts, c'est-à-dire, les processus dont les trajectoires peuvent avoir des discontinuités. Cependant, on leur imposera une certaine structure: pour définir la notion même d'un saut, il faut que les limites à gauche et à droite existent en tout point:

$$X_{t_-} := \lim_{s \uparrow t} X_s, \qquad X_{t_+} := \lim_{s \downarrow t} X_s$$

doivent exister pour tout t. Le saut au point t est alors défini par $\Delta X_t = X_{t+} - X_{t-}$. La valeur en t peut être égale à X_{t-} (processus càglàd pour continu à gauche, limite à droite) ou à X_{t+} (càdlàg pour continu à droite, limite à gauche). La plupart de processus dans ce chapitre seront des processus càdlàg.

5.1 Introduction aux processus de Lévy

Les processus de Lévy forment une classe de processus avec sauts qui est à la fois suffisamment simple pour étudier et en même temps assez riche pour les applications ou en tout cas pour être utilisé comme brique de base pour construire des modèles plus réalistes.

Définition 2. Un processus stochastique X est un processus de Lévy s'il est càdlàg, satisfait $X_0 = 0$ et possède les deux propriétés suivantes:

- Accroissements indépendants;
- Accroissements stationnaires;

De ces deux propriétés decoulent notamment

- Continuité en probabilité: $\forall \varepsilon, \lim_{s \to 0} P[|X_{s+t} - X_t| > \varepsilon] = 0$.
- Absence de sauts aux temps fixes: $\forall t$, $P[X_{t-} = X_t] = 1$.

Les processus de Lévy sont essentiellement des processus avec sauts, car on peut démontrer que tout processus de Lévy continu est un mouvement Brownien avec drift.

Proposition 3. *Soit X un processus de Lévy continu. Alors il existent $\gamma \in \mathbb{R}^d$ et une matrice symétrique definie positive A tels que*

$$X_t = \gamma t + W_t,$$

avec W un mouvement Brownien de matrice de variance-covariance A.

Preuve. Ce résultat est une conséquence directe du théorème de limite centrale de Feller-Lévy mais il est important pour la compréhension des processus de Lévy, on donnera donc une idée de preuve (dans le cas unidimensionnel).

Il suffit de demontrer que X_1 suit la loi gaussienne, le reste est une conséquence de l'indépendance et stationnarité des accroissements.

Etape 1 Soit $\xi_n^k := X_{\frac{k}{n}} - X_{\frac{k-1}{n}}$. La continuité de X implique

$$\lim_n P[\sup_k |\xi_n^k| > \varepsilon] = 0,$$

pour tout ε. Posons $a_n = P[|\xi_n^1| > \varepsilon]$. Puisque

$$P[\sup_k |\xi_n^k| > \varepsilon] = 1 - (1 - P[|\xi_n^1| > \varepsilon])^n,$$

on a $\lim_n (1-a_n)^n = 1$, d'où $\lim_n n\log(1-a_n) = 0$. Or $n\log(1-a_n) \le -na_n \le 0$. Nous avons donc

$$\lim_n nP[|X_{\frac{1}{n}}| > \varepsilon] = 0 \tag{5.1}$$

Etape 2 En utilisant l'indépendance et la stationnarité des accroissements, on démontre

$$\lim_n nE[\cos X_{\frac{1}{n}} - 1] = \frac{1}{2}\{\log Ee^{iX_1} + \log Ee^{-iX_1}\} := -A; \tag{5.2}$$

$$\lim_n nE[\sin X_{\frac{1}{n}}] = \frac{1}{2i}\{\log Ee^{iX_1} - \log Ee^{-iX_1}\} := \gamma. \tag{5.3}$$

Etape 3 Les équations (5.1) and (5.2) permettent de démontrer que pour toute fonction f telle que $f(x) = o(|x|^2)$ dans un voisinage de 0, $\lim_n nE[f(X_{\frac{1}{n}})] = 0$ ce qui implique pour tout $\varepsilon > 0$

$$\lim_n nE[X_{\frac{1}{n}} 1_{|X_{\frac{1}{n}}| \le \varepsilon}] = \gamma, \tag{5.4}$$

$$\lim_n nE[X_{\frac{1}{n}}^2 1_{|X_{\frac{1}{n}}| \le \varepsilon}] = A, \tag{5.5}$$

$$\lim_n nE[|X_{\frac{1}{n}}|^3 1_{|X_{\frac{1}{n}}| \le \varepsilon}] = 0. \tag{5.6}$$

$$\tag{5.7}$$

Etape 4 En rassemblant les différentes estimations, on a finalement

$$\begin{aligned}
\log E[e^{iuX_1}] &= n\log E[e^{iuX_{\frac{1}{n}}} 1_{X_{\frac{1}{n}} \le \varepsilon}] + o(1) \\
&= n\log\{1 + iuE[X_{\frac{1}{n}} 1_{X_{\frac{1}{n}} \le \varepsilon}] - \frac{u^2}{2} E[X_{\frac{1}{n}}^2 1_{X_{\frac{1}{n}} \le \varepsilon}] + o(1/n)\} + o(1) \\
&= iu\gamma - \frac{Au^2}{2} + o(1) \xrightarrow[n\to\infty]{} iu\gamma - \frac{Au^2}{2}
\end{aligned}$$

où $o(1)$ signifie une quantité qui tend vers 0 lorsque $n \to \infty$. □

L'autre exemple fondamental d'un processus de Lévy est le processus de Poisson.

Le processus de Poisson

Définition 3. Soit $(\tau_i)_{i\ge 1}$ une suite de v.a. exponentielles de paramètre λ et soit $T_n = \sum_{i=1}^n \tau_i$. Alors le processus

$$N_t = \sum_{n\ge 1} 1_{t \ge T_n} \tag{5.8}$$

est appellé le processus de Poisson de paramètre λ

Proposition 4 (Propriétés du processus de Poisson)**.**

1. *Pour tout $t \ge 0$, la somme dans* (5.8) *est finie p.s.*

2. *Les trajectoires de N sont constantes par morceaux avec des sauts de taille* 1 *seulement.*

3. *Les trajectoires sont càdlàg.*

4. *$\forall t > 0$, $N_{t-} = N_t$ avec probabilité* 1.

5. *$\forall t > 0$, N_t suit une loi de Poisson de paramètre λt:*

$$P[N_t = n] = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

6. *La fonction caractéristique du processus de Poisson est*

$$E[e^{iuN_t}] = \exp\{\lambda t(e^{iu} - 1)\}.$$

7. *Le processus de Poisson est un processus de Lévy.*

Le processus de Poisson compte les événements dont les temps d'attente sont des variables exponentielles indépendantes. Dans un cadre plus général, on parle d'un processus de comptage.

Définition 4. Soit (T_n) une suite de temps avec $T_n \to \infty$ p.s. Alors le processus

$$N_t = \sum_{n\geq 1} 1_{t\geq T_n}$$

est appellé un processus de comptage.

Autrement dit, un processus de comptage est un processus croissant constant par morceaux, avec sauts de taille 1 et p.s. fini.

Un premier pas vers la caractérisation d'un processus de Lévy est de caractériser les processus de Lévy de comptage.

Proposition 5. *Soit (N_t) un processus de Lévy et un processus de comptage. Alors (N_t) est un processus de Poisson.*

Preuve. La preuve repose sur la caractérisation de la distribution exponentielle par la propriété d'absence de memoire: si pour une variable aléatoire T on a

$$P[T > t + s|T > t] = P[T > s]$$

pour tout $t, s > 0$ alors T est une v.a. exponentielle.

Soit T_1 le 1er temps de saut du processus N. L'indépendance et la stationnarité des accroissements donnent:

$$\begin{aligned} P[T_1 > t + s|T_1 > t] &= P[N_{t+s} = 0|N_t = 0] \\ &= P[N_{t+s} - N_t = 0|N_t = 0] = P[N_s = 0] = P[T_1 > s], \end{aligned}$$

le premier temps de saut T_1 est donc une variable exponentielle.

Maintenant, il suffit de demontrer que le processus $(X_{T_1+t} - X_{T_1})_{t\geq 0}$ est indépendant du X_{T_1} et a la même loi que $(X_t)_{t\geq 0}$. Soit $f(t) := E[e^{iuX_t}]$. Alors l'indépendance et la stationnarité des accroissements entrainent que $f(t+s) = f(t)f(s)$ et $M_t := \frac{e^{iuX_t}}{f(t)}$ est une martingale. Soit $T_1^n := n \wedge T_1$. Alors par application du théorème d'arrêt de Doob,

$$E[e^{iu(X_{T_1^n+t} - X_{T_1^n}) + ivT_1^n}] \quad = \quad E\left[\frac{f(T_1^n + t)}{f(T_1^n)}e^{ivT_1^n}\right] \quad = \quad E[e^{iuX_t}]E[e^{ivT_1^n}].$$

Pour terminer la preuve, il reste alors à appliquer le théorème de convergence dominée. □

Processus de Poisson composé Le processus de Poisson en lui-même n'est pas utilisé pour modéliser les cours d'actifs, car la condition que la taille de sauts est toujours égale à 1 est trop contraignante, mais il sert comme brique de base pour construire des modèles plus riches.

Définition 5 (Processus de Poisson composé)**.** Le processus de Poisson composé avec intensité des sauts λ et distribution de la taille des sauts μ est un processus stochastique $(X_t)_{t\geq 0}$ défini par

$$X_t = \sum_{i=1}^{N_t} Y_i,$$

où $\{Y_i\}_{i\geq 1}$ est une suite de v.a. indépendantes de loi μ et N est un processus de Poisson standard d'intensité λ indépendant de $\{Y_i\}_{i\geq 1}$.

En d'autres mots, un processus de Poisson composé est un processus constant par morceaux qui saute aux instants de sauts d'un processus de Poisson standard, et dont les tailles de sauts sont des variables i.i.d. d'une loi donnée.

Proposition 6 (Propriétés du processus de Poisson composé)**.** *Soit $(X_t)_{t\geq 0}$ un processus de Poisson composé d'intensité des sauts λ et de loi des sauts μ. Alors X est un processus de Lévy constant par morceaux et sa fonction caractéristique est donnée par*

$$E[e^{iuX_t}] = \exp\left\{t\lambda \int_{\mathbb{R}} (e^{iux} - 1)\mu(dx)\right\}.$$

Exemple 5 (Modèle de Merton)**.** Le modèle de Merton (1976) est l'une des premières applications de processus avec sauts en modélisation financière. Dans ce modèle, pour prendre en compte les discontinuités dans les cours d'actions, on rajoute des sauts gaussiens au logarithme du prix:

$$S_t = S_0 e^{rt+X_t}, \quad X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_i, \quad Y_i \sim N(\mu, \delta^2) \text{ indépendants.}$$

L'avantage de cette modélisation est d'avoir une representation en série pour la densité de probabilité du log-prix (et pour les prix d'options européennes):

$$p_t(x) = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k \exp\left\{-\frac{(x-\gamma t - k\mu)^2}{2(\sigma^2 t + k\delta^2)}\right\}}{k!\sqrt{2\pi(\sigma^2 t + k\delta^2)}}.$$

Mesures aléatoires de Poisson La notion de mesure aléatoire de Poisson est centrale pour toute la théorie; elle nous permettra de donner la description complète des trajectoires d'un processus de Lévy.

Définition 6 (Mesure aléatoire)**.** Soit $(\Omega, P, \mathcal{F})$ un espace de probabilité et $(E, \mathcal{E})$ un espace mesurable. Alors $M : \Omega \times \mathcal{E} \to \mathbb{R}$ est une mesure aléatoire si

- Pour chaque $\omega \in \Omega$, $M(\omega, \cdot)$ est une mesure sur $\mathcal{E}$.
- Pour chaque $A \in \mathcal{E}$, $M(\cdot, A)$ est mesurable.

Définition 7 (Mesure aléatoire de Poisson)**.** Soit $(\Omega, P, \mathcal{F})$ un espace de probabilité, $(E, \mathcal{E})$ un espace mesurable et μ une mesure sur $(E, \mathcal{E})$. Alors $M : \Omega \times \mathcal{E} \to \mathbb{R}$ est une mesure aléatoire de Poisson d'intensité μ si

- Pour tout $A \in \mathcal{E}$ avec $\mu(A) < \infty$, $M(A)$ suit la loi de Poisson d'intensité $E[M(A)] = \mu(A)$.
- Pour tous $A_1, \dots A_n$ disjoints, $M(A_1), \dots, M(A_n)$ sont indépendants.

En particulier, la mesure aléatoire de Poisson est une mesure aléatoire positive aux valeurs entières. Elle peut être construite comme la mesure de comptage d'un nuage de points comme le montre la proposition suivante.

Proposition 7. *Soit μ une mesure σ-finie sur un sous-ensemble mesurable E de $\mathbb{R}^d$. Alors il existe une mesure de Poisson sur E d'intensité μ.*

Preuve. Supposons dans un premier temps que $\mu(E) < \infty$. Soit $\{X_i\}_{i\geq 1}$ une suite de v.a. indépendantes telles que $P[X_i \in A] = \frac{\mu(A)}{\mu(E)}, \forall i$ et $\forall A \in \mathcal{B}(E)$, et soit $M(E)$ une v.a. de Poisson d'intensité $\mu(E)$ indépendante de $\{X_i\}_{i\geq 1}$. Il est alors facile de demontrer que la mesure aléatoire M définie par

$$M(A) := \sum_{i=1}^{M(E)} 1_A(X_i), \quad \forall A \in \mathcal{B}(E),$$

est une mesure aléatoire de Poisson sur E d'intensité μ.

Si maintenant $\mu(E) = \infty$, il suffit de prendre une suite d'ensembles mesurables disjoints $\{E_i\}_{i\geq 1}$ telle que $\mu(E_i) < \infty, \forall i$ et $\bigcup_i E_i = E$, construire une mesure de Poisson M_i sur chaque E_i avec la procedure décrite ci-dessus et poser

$$M(A) := \sum_{i=1}^{\infty} M_i(A), \quad \forall A \in \mathcal{B}(E).$$

□

Corollaire 1 (Formule exponentielle)**.** *Soit M une mesure aléatoire de Poisson sur $(E, \mathcal{E})$ d'intensité μ, $B \in \mathcal{E}$ et f une fonction mesurable avec $\int_B |e^{f(x)} - 1|\mu(dx) < \infty$. Alors*

$$E\left[e^{\int_B f(x)M(dx)}\right] = \exp\left[\int_B (e^{f(x)} - 1)\mu(dx)\right].$$

Définition 8 (Mesure de sauts)**.** Soit X un processus càdlàg aux valeurs dans $\mathbb{R}^d$. La mesure de sauts de X est une mesure aléatoire sur $\mathcal{B}([0,\infty) \times \mathbb{R}^d)$ définie par

$$J_X(A) = \#\{t : \Delta X_t \neq 0 \text{ and } (t, \Delta X_t) \in A\}.$$

La mesure de sauts d'un ensemble de type $[s,t] \times A$ compte le nombre de sauts de X entre s et t dont les tailles tombent dans A. Pour un processus de comptage, comme la taille de saut est toujours égale à 1, la mesure de sauts est une mesure aléatoire sur $[0,\infty)$ simplement.

Proposition 8. *Soit X un processus de Poisson d'intensité λ. Alors J_X est une mesure aléatoire de Poisson sur $[0,\infty)$ d'intensité $\lambda \times dt$.*

Le résultat peut-être le plus important de la théorie de processus de Lévy est que la mesure de sauts d'un processus de Lévy général est également une mesure aléatoire de Poisson.

Exercice 5. Soit X et Y deux processus de Lévy indépendants. A partir de la définition, démontrer que $X+Y$ est un processus de Lévy.

Exercice 6. Demontrer que la propriété d'absence de mémoire caractérise la loi exponentielle: si une variable aléatoire T satisfait

$$\forall t, s > 0, \quad P[T > t + s | T > t] = P[T > s]$$

alors soit $T \equiv 0$ soit T suit la loi exponentielle.

Exercice 7. Démontrer la propriété 7 du processus de Poisson (si N est un processus de Poisson alors il est un processus de Lévy).

Exercice 8. Demontrer que si N et N' sont deux processus de Poisson indépendants de paramètres λ et λ' alors $N + N'$ est un processus de Poisson de paramètre $\lambda + \lambda'$.

Exercice 9. Soit X un processus de Poisson composé, avec la loi de sauts μ. Etablir que

- $E[|X_t|] < \infty$ si et seulement si $\int_{\mathbb{R}} |x| f(dx)$ et dans ce cas

$$E[X_t] = \lambda t \int_{\mathbb{R}} x f(dx).$$

- $E[|X_t|^2] < \infty$ si et seulement si $\int_{\mathbb{R}} x^2 f(dx)$ et dans ce cas

$$\mathrm{Var}[X_t] = \lambda t \int_{\mathbb{R}} x^2 f(dx).$$

- $E[e^{X_t}] < \infty$ si et seulement si $\int_{\mathbb{R}} e^x f(dx)$ et dans ce cas

$$E[e^{X_t}] = \exp\left(\lambda t \int_{\mathbb{R}} (e^x - 1) f(dx)\right).$$

Exercice 10. Ici il s'agit de demontrer que pour construire une mesure de Poisson sur $\mathbb{R}$, il faut prendre deux processus de Poisson et faire partir un vers $+\infty$ et l'autre vers $-\infty$.

Soit N et N' deux processus de Poisson d'intensité λ, et soit M une mesure aléatoire définie par

$$M(A) = \#\{t > 0 : t \in A, \Delta N_t = 1\} + \#\{t > 0 : -t \in A, \Delta N'_t = 1\}.$$

Montrer que M est une mesure aléatoire de Poisson d'intensité λ.

5.2 Structure des trajectoires d'un processus de Lévy

Définition 9 (Mesure de Lévy)**.** Soit X un processus de Lévy à valeurs dans $\mathbb{R}^d$. Alors la mesure ν définie par

$$\nu(A) = E[\#\{t \in [0,1] : \Delta X_t \neq 0 \text{ et } \Delta X_t \in A\}],\ A \in \mathcal{B}(\mathbb{R}^d)$$

est appellée la mesure de Lévy de X.

Théorème 3 (Décomposition de Lévy-Itô)**.** *Soit X un processus de Lévy à valeurs dans $\mathbb{R}^d$ de mesure de Lévy ν. Alors*

1. *Sa mesure de sauts J_X est une mesure aléatoire de Poisson sur $[0,\infty) \times \mathbb{R}^d$ d'intensité $dt \times \nu$.*
2. *Sa mesure de Lévy satisfait, $\int_{\mathbb{R}^d} (\|x\|^2 \wedge 1)\nu(dx) < \infty$.*
3. *Il existent $\gamma \in \mathbb{R}^d$ et un mouvement Brownien d-dimensionnel B de matrice de covariance A tels que*

$$\begin{aligned} X_t &= \gamma t + B_t + N_t + M_t, \quad \text{où} \\ N_t &= \int_{|x|>1, s\in[0,t]} x J_X(ds \times dx) \quad \text{and} \\ M_t &= \int_{0<|x|\leq 1, s\in[0,t]} x\{J_X(ds \times dx) - \nu(dx)ds\} \\ &\equiv \int_{0<|x|\leq 1, s\in[0,t]} x \tilde{J}_X(ds \times dx). \end{aligned} \tag{5.9}$$

Les trois termes sont indépendants et la convergence dans le dernier terme est presque sure et uniforme en t sur les compactes.

Le triplet (A, ν, γ) est appellé le *triplet caractéristique* du processus X. La demonstration est fondée sur le lemme suivant:

Lemme 1. *Soit (X, Y) un processus de Lévy 2-dimensionnel tel que Y est constant par morceaux et $\Delta X_t \Delta Y_t = 0$ pour tout t p.s. Alors X et Y sont indépendants.*

Preuve. En vue de l'indépendance et la stationnarité des accroissements, il suffit de demontrer que X_1 et Y_1 sont indépendants. Soit $M_t = \frac{e^{iuX_t}}{E[e^{iuX_t}]}$ et $N_t = \frac{e^{iuY_t}}{E[e^{iuY_t}]}$. Alors M et N sont martingales sur $[0, 1]$. De l'indépendance et stationnarité des accroissements on déduit que pour tout processus de Lévy Z,

$$E[e^{iuZ_t}] = E[e^{iuZ_1}]^t \quad \text{et} \quad E[e^{iuZ_1}] \neq 0, \forall u.$$

Ceci implique que M est borné. Par proposition 5, le nombre de sauts de Y sur $[0, 1]$ est une variable de Poisson. N est donc à variation intégrable sur cette intervalle. Par la propriété de martingale et convergence dominée on a finalement

$$E[M_1 N_1] - 1 = E\left[\sum_{i=1}^{n} (M_{i/n} - M_{(i-1)/n})(N_{i/n} - N_{(i-1)/n})\right] \to E\left[\sum_{0 \leq t \leq 1} \Delta M_t \Delta N_t\right] = 0,$$

ce qui implique $E[e^{iuX_1 + ivY_1}] = E[e^{iuX_1}]E[e^{ivY_1}]$. □

Preuve du théorème 3.

Partie 1 Soit $A \in \mathcal{B}(\mathbb{R}^d)$ avec $0 \notin \bar{A}$. Alors $N_t^A = \#\{s \leq t : \Delta X_s \in A\}$ est un processus de comptage et un processus de Lévy, donc, par proposition 5, un processus de Poisson, ce qui implique que $J_X([t_1, t_2] \times A)$ suit la loi de Poisson d'intensité $(t_2 - t_1)\nu(A)$ et que $J_X([t_1, t_2] \times A)$ est indépendant de $J_X([s_1, s_2] \times A)$ si $t_2 \leq s_1$. Prenons maintenant deux ensembles A et B disjoints. Par lemme 1, N^A et N^B sont indépendants, ce qui montre que $J_X([s_1, s_2] \times A)$ et $J_X([t_1, t_2] \times B)$ le sont aussi, pour tous s_1, s_2, t_1, t_2.

Partie 2 De la partie précedente, on déduit que $\nu(A) < \infty$ dès que $0 \notin \bar{A}$. Il reste à demontrer que

$$\int_{\|x\| \leq \delta} \|x\|^2 \nu(dx) < \infty$$

pour un $\delta > 0$. Soit

$$X_t^\varepsilon = \sum_{0 \leq s \leq t}^{1 \geq |\Delta X_s| > \epsilon} \Delta X_s = \int_{\varepsilon < |x| \leq 1, s \in [0,t]} x J_X(ds \times dx)$$

et $R_t^\varepsilon = X_t - X_t^\varepsilon$. Comme $(X_t^\varepsilon, R_t^\varepsilon)$ est un processus de Lévy, lemme 1 implique que X_t^ε et R_t^ε sont indépendants. De plus, $E[e^{i\langle u, X_t \rangle}] > 0$ pour tout t, u. Ceci implique

$$E[e^{iuX_t}] = E[e^{iuR_t^\varepsilon}]E[e^{iuX_t^\varepsilon}].$$

Donc, $E[e^{iuX_t^\varepsilon}]$ est borné inférieurement par un nombre positif indépendant de ε. Par la formule exponentielle, ceci equivaut à

$$\left|\exp\left\{t \int_{|x| \geq \varepsilon} (e^{iux} - 1)\nu(dx)\right\}\right| \geq C > 0,$$

ce qui implique $\int_{|x| \geq \varepsilon} (1 - \cos(ux))\nu(dx) \leq \tilde{C} < \infty$. Puisque ce résultat est vrai pour tout u, la preuve de la partie 2 est terminée.

Partie 3 Notons tout d'abord que le processus M est bien défini grâce à la compensation des petits sauts et au fait que la mesure de Lévy integre $\|x\|^2$ pres de zéro: en introduisant le processus

$$M_t^\varepsilon = \int_{\varepsilon \leq \|x\| < 1, s \in [0,t]} x \tilde{J}_X(ds \times dx),$$

on obtient que pour $\varepsilon_1 < \varepsilon_2$,

$$E[(M_t^{\varepsilon_1} - M_t^{\varepsilon_2})^2] = t \int_{\varepsilon_2 \le \|x\| < \varepsilon_1} x^2 \nu(dx)$$

et donc, puisque l'espace L^2 est complet, que pour chaque t, M_t^ε converge en L^2 lorsque $\varepsilon \to 0$. En appliquant l'inégalité de Doob, on montre que la convergence est uniforme en t sur les intervalles compactes. Le processus $X - N - M$ est donc un processus de Lévy continu indépendant de N et M en vertu du lemme 1. On conclut en appliquant proposition 3. □

Corollaire 2 (Lévy-Khintchine representation)**.** *Soit X un processus de Lévy de triplet caractéristique (A, ν, γ). Sa fonction caractéristique est donnée par*

$$E[e^{iuX_t}] = \exp\left\{ t \left(i\gamma u - \frac{Au^2}{2} + \int_{\mathbb{R}} (e^{iux} - 1 - iux 1_{|x|\le 1}) \nu(dx) \right) \right\}, \tag{5.10}$$

Preuve. En utilisant le théorème précedent et la formule exponentielle, on a

$$E[e^{iu(\gamma t + B_t + N_t + M_t^\varepsilon)}] = \exp\left\{ t \left(i\gamma u - \frac{Au^2}{2} + \int_{|x|\ge\varepsilon} (e^{iux} - 1 - iux 1_{|x|\le 1}) \nu(dx) \right) \right\},$$

et le théorème de convergence dominée permet de conclure. □

Le processus variance gamma L'un des exemples les plus simples de processus de Lévy avec intensité infinie de sauts est le processus gamma, un processus aux accroissements indépendants et stationnaires tel que pour tout t, la loi p_t de X_t est la loi gamma de paramètres λ et ct:

$$p_t(x) = \frac{\lambda^{ct}}{\Gamma(ct)} x^{ct-1} e^{-\lambda x}.$$

Le processus gamma est un processus de Lévy croissant dont la fonction caractèristique a une forme très simple

$$E[e^{iuX_t}] = (1 - iu/\lambda)^{-ct}.$$

On démontre facilement que la mesure de Lévy du processus gamma a une densité donnée par

$$\nu(x) = \frac{ce^{-\lambda x}}{x} 1_{x>0}. \tag{5.11}$$

A partir du processus gamma, on peut construire un modèle avec sauts très populaire: le processus variance gamma [49, 47], qui est obtenu en changeant l'echelle de temps d'un mouvement brownien avec drift par un processus gamma:

$$Y_t = \mu X_t + \sigma B_{X_t}.$$

L'utilisation de Y_t pour modéliser le logarithme du prix d'action est habituellement justifiée en disant que le prix suit un mouvement Brownien géometrique sur une échelle de temps stochastique donnée par le processus gamma [34]. Le processus variance gamma est un autre exemple du processus de Lévy avec intensité infinie de sauts, et sa fonction caractéristique est donnée par

$$E[e^{iuY_t}] = \left(1 + \frac{\kappa\sigma^2 u^2}{2} - i\mu\kappa u\right)^{-\kappa t}.$$

Le paramètres ont l'interprétation intuitive suivante: σ est un paramètre d'échelle, μ est le paramètre d'asymétrie (skewness) et κ est responsable pour le kurtosis du processus (épaisseur des queues de la densité).

Exercice 11. Soit X un processus de Lévy avec triplet caractéristique (A, ν, γ). Calculez la probabilité que X aura au moins un saut négatif de taille supérieure à $\varepsilon > 0$ sur l'intervalle $[0, T]$.

Exercice 12. Soit X un processus de Lévy avec la mesure de Lévy de la forme $\nu(dx) = \lambda\nu_0(dx)$, où ν_0 n'a pas d'atomes et satisfait $\nu_0(\mathbb{R}) = \infty$. Pour tout $n \in \mathbb{N}$, $k_n > 0$ est la solution de $\int_{k_n}^{\infty} \nu_0(dx) = n$. Pour T fixé, quelle est la loi de la variable aléatoire

$$A_n = \#\{t \leq T : \Delta X_t \in [k_{n+1}, k_n]\}.$$

Utiliser ce résultat pour proposer une méthode d'estimation de λ à partir de l'observation d'une trajectoire de X, en supposant que ν_0 est connu.

Exercice 13. Soit X un processus de Lévy avec sans composante de diffusion et avec une mesure de Lévy ν qui satisfait $\int_{\mathbb{R}} |x|\nu(dx) < \infty$. A l'aide de la décomposition de Lévy-Itô, montrer que les trajectoires de X sont p.s. à variation finie (on dit qu'une fonction est à variation finie si elle peut être representée comme différence de deux fonctions croissantes).

Exercice 14. Demontrer que la mesure de Lévy du processus gamma est donnée par l'équation (5.11). Montrer que le processus variance gamma peut être représenté comme une différence de deux processus gamma indépendants, et en deduire la forme de la mesure de Lévy du processus variance gamma.

5.3 Modèles exponentielle-Lévy

Pour assurer la positivité des prix, on modélise souvent les cours d'actions comme exponentielles des processus de Lévy:

$$S_t = S_0 e^{rt + X_t}. \tag{5.12}$$

Le modèle (5.12) n'admet pas d'opportunité d'arbitrage s'il existe une probabilité historique Q équivalente à P telle que e^X est une Q-martingale. Si le triplet caractéristique de X sous Q est (A^Q, ν^Q, γ^Q), alors la condition de martingale s'écrit

$$\gamma + \frac{A}{2} + \int_{\mathbb{R}} (e^x - 1 - x1_{|x|\leq 1})\nu^Q(dx) = 0.$$

Pour appliquer les modèles (5.12) à l'évaluation d'options, il faut donc établir l'existence d'une probabilité martingale équivalente, et pour cela nous avons besoin de répondre à la question suivante:

- Etant donnés 2 processus de Lévy (X, P) et (X, Q), avec triplets caractéristiques (A^P, ν^P, γ^P) et (A^Q, ν^Q, γ^Q), comment déterminer si les probabilités P et Q sont équivalentes?

Dans le cas continu (si $\nu^P \equiv \nu^Q \equiv 0$), on sait que la condition nécessaire et suffisante est $A^P = A^Q > 0$. Pour les processus de Poisson composés, la solution est facile:

Proposition 9. *Soit (X, P) et (X, Q) deux processus de Poisson composés de mesures de Lévy ν^P et ν^Q. Alors $P \sim Q$ sur $[0, T]$ si et seulement si $\nu^P \sim \nu^Q$ et dans ce cas,*

$$\frac{dQ}{dP}|_{\mathcal{F}_T} = \exp\left(T(\lambda^P - \lambda^Q) + \sum_{s\leq T:\Delta X_s \neq 0} \phi(\Delta X_s)\right) := \exp(Z_T), \tag{5.13}$$

où $\lambda^P := \nu^P(\mathbb{R})$, $\lambda^Q = \nu^Q(\mathbb{R})$ et $\phi = \log \frac{d\nu^Q}{d\nu^P}$.

Preuve. Supposons dans un premier temps que $\nu^Q \sim \nu^P$. Il est clair alors que Z est un processus de Lévy dont l'intensité de sauts est égale à λ^P. De plus,

$$T(\lambda^P - \lambda^Q) + T\int \nu^Z(dx)(e^x - 1) = 0$$

montre que e^Z est une martingale et donc un changement de probabilité. Soit une probabilité Q^* équivalente à P définie par

$$\frac{dQ^*}{dP}|_{\mathcal{F}_T} = \exp(Z_T).$$

Sois la probabilité Q, X est un processus de Lévy puisque Z l'est sous P. On calcule la fonction caractéristique de X sous Q^*:

$$E^{Q^*}[e^{iuX_t}] = E^P[e^{iuX_t}e^{Z_t}] = e^{t(\lambda^P-\lambda^Q)}\exp\left(t\int(e^{iu(x+\phi(x))}-1)\nu^P(dx)\right)$$
$$= \exp\left(t\int(e^{iux}-1)\nu^Q(dx)\right).$$

Donc, la mesure de Lévy de X sous Q^* est égale à ν^Q, ce qui implique que Q^* coïncide avec Q et donc $Q \sim P$.

Si, au contraire, ν^Q n'est pas équivalent à ν^P, il existe $A \subset \mathbb{R}$ tel que soit $\nu^Q(A) > 0$ et $\nu^P(A) = 0$ soit $\nu^Q(A) = 0$ et $\nu^P(A) > 0$. Supposons pour fixer les idées qu'on est dans le premier cas. Alors l'evenement "il y a au moins un saut de X dans $[0,T]$ dont la taille tombe dans A" a probabilité nulle sous P et non-nulle sous Q, ce qui veut dire que Q n'est pas equivalente à P. □

Le résultat précedent montre que dans le contexte de processus de Poisson on peut changer librement l'intensité de chaque taille de saut, mais on ne peut pas completement enlever les sauts d'une taille donnée ni rajouter des sauts d'une taille qui était absente sous la probabilité originale. Contrairement à ce qui se passe dans le cas du mouvement brownien, on ne peut pas non plus changer le drift du processus, car il peut être observé de manière presque sure à partir d'une seule trajectoire.

Lorsque l'intensité de sauts est infinie, exercice 12 suggère que des contraintes supplémentaires sur la taille des petits sauts sont à respecter. Le théorème suivant de [55], reproduit ici sans preuve, donne des conditions nécessaires et suffisantes.

Théorème 4. *Soient (X,P) et (X,Q) processus de Lévy avec triplets caractéristiques (A^P,ν^P,γ^P) et (A^Q,ν^Q,γ^Q). Alors $P \sim Q$ sur $[0,T]$ pour tout T si et seulement si*

1. *Les coefficients de diffusion verifient $A^Q = A^P := A$.*
2. *Les mesures de Lévy verifient $\nu^Q \sim \nu^P$ avec la fonction $\phi := \log\frac{d\nu^Q}{d\nu^P}$ qui satisfait*
$$\int(e^{\phi/2}-1)^2\nu^P(dx) < \infty.$$
3. *Il existe $\beta \in \mathbb{R}$ avec*
$$\gamma^Q = \gamma^P + \int_{|x|\le 1}(e^\phi - 1)\nu^P(dx) + \beta A. \tag{5.14}$$

Cette dernière condition est une contrainte seulement si $A = 0$.

On peut enfin analyser l'existence d'opportunités d'arbitrage dans les modèles exponentielle-Lévy. La proposition suivante montre que ces modèles ne sont "presque jamais" arbitrables.

Proposition 10. *Soit (X,P) un processus de Lévy de triplet caractéristique (A,ν,γ). Il existe une probabilité $Q \sim P$ telle que e^X est une Q-martingale si et seulement si X n'est pas p.s. croissant ni p.s. decroissant.*

Preuve. La partie "seulement si" est quasi-triviale, on se concentre sur la partie "si".

Si $A > 0$, on peut obtenir une probabilité équivalente par un simple changement de drift. On suppose donc sans perte de généralité que $A = 0$. De plus on peut supposer que $\int_{|x|\ge 1} e^{\theta x}\nu(dx) < \infty$ pour tout θ car sinon on peut commencer par un changement de probabilité équivalente avec $\nu^Q = e^{-x^2}\nu$.

Pour un $\theta \in \mathbb{R}$, soit $\nu^Q = e^{\theta x}\nu$ et $\gamma^Q = \gamma + \int_{|x|\le 1}(e^{\theta x}-1)\nu(dx)$ et soit (X,Q) un processus de Lévy de triplet caractéristique $(0,\nu^Q,\gamma^Q)$. Alors par le théorème précedent, $Q \sim P$. Pour demontrer qu'il existe une probabilité martingale équivalente, il faut trouver un θ tel que

$$\gamma^Q + \int(e^x - 1 - x1_{|x|\le 1})\nu^Q(dx) = 0,$$

ce qui est équivalent à

$$f(\theta) := \gamma + \int (e^{\theta(x+1)} - e^{\theta x} - x1_{x\leq 1})\nu(dx) = 0.$$

La fonction $f(\theta)$ est croissante de dérivée

$$f'(\theta) = \int xe^{\theta x}(e^x - 1)\nu(dx) \geq 0.$$

Si $\nu((-\infty, 0)) > 0$ et $\nu((0, \infty)) > 0$, la dérivée f' est bornée inférieurement par

$$\min\left(\int_{-\infty}^{0} x(e^x - 1)\nu(dx), \int_0^{\infty} x(e^x - 1)\nu(dx)\right)$$

ce qui implique que dans ce cas $f(\theta) = 0$ a une solution. Supposons donc pour fixer les idées que $\nu((-\infty, 0)) = 0$. Dans ce cas

$$\lim_{\theta\to+\infty} f(\theta) = +\infty$$

et

$$\lim_{\theta\to-\infty} f(\theta) = \lim_{\theta\to-\infty} \{\gamma + \int_0^1 x(e^{\theta x} - 1)\}.$$

Lorsque $\int_0^1 x\nu(dx) = \infty$ (processus à variation infinie), cette limite vaut $-\infty$ et donc $f(\theta) = 0$ a une solution. Dans le cas contraire, cette limite vaut

$$\gamma - \int_0^1 x\nu(dx),$$

et $f(\theta) = 0$ a une solution seulement si $\gamma - \int_0^1 x\nu(dx) < 0$, c'est-à-dire, le drift du processus doit être négatif. Comme on a supposé que les sauts sont positifs, ceci implique que X ne doit pas être croissant. De la même manière, on montre qu'il ne doit pas non plus être décroissant. □

Exercice 15. Soit (X, P) un processus de Lévy avec triplet caractéristique (A, ν, γ), et soit Q une mesure de probabilité définie par

$$\frac{dQ}{dP}|_{\mathcal{F}_t} = \frac{e^{\theta X_t}}{E[e^{\theta X_t}]}$$

pour un $\theta \in \mathbb{R}$ tel que $E[e^{\theta X_t}] < \infty$. Calculer le triplet caractéristique de X sous Q.

Exercice 16.

- Soit X un processus de Lévy et soit $f : \mathbb{R} \to (0, \infty)$. Quelle condition doit-on imposer sur la fonction f pour que la somme
$$\sum_{t\in[0,1]:\Delta X_t\neq 0} f(\Delta X_t)$$
soit convergente p.s.?
- Soit (X, P) un processus de Lévy de mesure de Lévy $\frac{1}{|x|^{1+\alpha}}$ et (X, Q) un processus de Lévy de mesure de Lévy $\frac{1}{|x|^{1+\alpha'}}$ (avec $\alpha > 0$ et $\alpha' > 0$). Montrer, en utilisant la question précedente, que $P \sim Q$ implique $\alpha = \alpha'$. Verifier votre résultat à l'aide du théorème 4.

5.4 Bases du calcul stochastique pour les processus avec sauts

Integrands et integrateurs L'application principale de l'integrale stochastique en finance est la représentation d'un portefeuille autofinançant: en l'absence de taux d'intérêt, lorsque le prix de l'actif risqué est un processus aux trajectoires continues S et la quantité de l'actif est ϕ, la valeur du portefeuille est

$$V_T = \int_0^T \phi_t dS_t$$

On voudrait que cette relation soit aussi vraie en présence de sauts, mais quelle sont les propriétés naturelles à imposer sur S et ϕ? Le processus S doit être continu à droite car les sauts dans les prix arrivent de façon inattendue. La stratégie de couverture ϕ_t, au contraire, est fondé sur les observations du gerant du portefeuille avant date t; elle doit donc être continue à gauche. Une illustration est donnée par l'exemple suivant: supposons que le prix d'un actif est décrit par $S_t = \lambda t - N_t$, où N_t est un processus de Poisson d'intensité λ, et soit T le temps du premier saut de N. Si on pouvait choisir la stratégie càdlàg $\phi_t = 1_{[0,T)}(t)$, consistant a vendre l'actif juste avant le saut, on aurait une opportunité d'arbitrage car

$$V_t = \int_0^t \phi_t dS_t = \lambda t \wedge T.$$

Avec la stratégie càglàd $\phi_t = 1_{[0,T]}(t)$, on a

$$V_t = \int_0^t \phi_t dS_t = \lambda t \wedge T - N_{t\wedge T},$$

ce qui est d'espérance nulle. Il est donc naturel de considerer les intégrands qui sont adaptés et continus à gauche. La forme la plus simple (et la seule réalisable en pratique) d'une stratégie de portefeuille est celle où le portefeuille n'est rebalancé qu'un nombre fini de fois. On définit alors un processus prévisible simple par

$$\phi_t = \phi_0 1_{t=0} + \sum_{i=0}^{n} \phi_i 1_{(T_i, T_{i+1}]}(t), \tag{5.15}$$

où $T_0 = 0$, $(T_i)_{i\geq 0}$ est une suite de temps d'arrêt et pour chaque i, ϕ_i est $\mathcal{F}_{T_i}$-mesurable et borné. L'espace de processus prévisibles simples sera noté par $\mathcal{S}$.

Pour les processus prévisibles simples, on définit l'intégrale stochastique par

$$\int_0^t \phi_s dS_s := \sum_{i=0}^{n} \phi_i (S_{T_{i+1}\wedge t} - S_{T_i \wedge t}) \tag{5.16}$$

Pour un processus général adapté et continu à gauche, l'intégrale stochastique est défini avec une extension par continuité, en utilisant la topologie de convergence en probabilité uniformement sur le compactes (ucp): on dit que la suite (X^n) de processus converge vers X en probabilité uniformement sur le compactes si, pour tout t, $(X^n - X)^*_t$ converge vers 0 en probabilité, où $Z^*_t := \sup_{0\leq s\leq t} |Z_s|$. On note par $\mathcal{S}_{ucp}$ l'espace $\mathcal{S}$ avec la topologie de convergence ucp, et par $\mathbb{L}_{ucp}$ et $\mathbb{D}_{ucp}$ l'espace de processus adaptés continus, respectivement, à gauche et à droite, avec cette même topologie. Il est alors possible de démontrer que l'espace $\mathcal{S}_{ucp}$ est dense dans $\mathbb{L}_{ucp}$, et d'associer à la topologie ucp une métrique sur $\mathbb{D}_{ucp}$, pour laquelle cette espace sera complet. Pour effectuer une extension par continuité de l'opérateur d'intégration stochastique défini par (5.16) de $\mathcal{S}_{ucp}$ vers $\mathbb{L}_{ucp}$, il faut que cet opérateur soit un opérateur continu de $\mathcal{S}_{ucp}$ vers $\mathbb{D}_{ucp}$. Or, cela dépend de l'intégrateur S, et on va se limiter justement aux intégrateurs pour lesquels ceci est vrai.

Définition 10. Un processus $S \in \mathbb{D}$ est une *semimartingale* si l'opérateur d'intégration stochastique défini par (5.16) est un opérateur continu de $\mathcal{S}_{ucp}$ vers $\mathbb{D}_{ucp}$.

Tout processus càdlàg adapté et de variation finie sur les compactes est une sémimartingale. Ceci est une conséquence de

$$\sup_{0\leq t\leq T} \left| \int_0^t \phi_s dS_s \right| \leq \mathrm{Var}_0^T(S) \sup_{0\leq t\leq T} \phi_t,$$

où $\mathrm{Var}_0^T(S)$ signifie la variation totale de S sur $[0,T]$. Une martingale càdlàg de carrée intégrable est une sémimartingale. En effet, pour un processus prévisible simple ϕ de la forme (5.15), et une martingale de carrée intégrable M, $\int_0^\cdot \phi_t dM_t$ est également une martingale et

$$E\left(\int_0^T \phi_t dM_t \right)^2 \leq \sup_{0\leq t\leq T, \omega\in\Omega} \phi_t^2 E[M_T^2].$$

Soit maintenant (ϕ^n) une suite de processus prévisibles simples telle que $\phi^n \to 0$ in ucp. Alors, en utilisant l'inégalité de Chebychev et l'inégalité de Doob, on obtient

$$P\left[\left(\int_0^\cdot \phi_t^n dM_t\right)^* > \varepsilon\right] \leq P\left[\left(\int_0^\cdot \phi_t^n 1_{|\phi_t^n|\leq C} dM_t\right)^* > \varepsilon\right] + P[(\phi^n)^* > C]$$
$$\leq \frac{4C^2 E[M_T^2]}{\varepsilon^2} + P[(\phi^n)^* > C] \to 0,$$

parce que qu'on peut rendre le premier terme arbitrairement petit en choisissant C petit et le deuxième en choisissant n grand.

Puisque les termes γt et N_t dans la décomposition de Lévy-Itô sont de variation finie, et les termes B_t et M_t sont des martingales de carrée intégrable, on conclût que tout processus de Lévy est une sémimartingale.

Un résultat très profond de la théorie des processus [52] est que toute sémimartingale est la somme d'un processus de variation finie et une martingale *locale*. La notion de martingale locale étend la notion de martingale: un processus (X_t) est une martingale locale s'il existe une suite de temps d'arrêt $\{T_i\}_{i\geq 1}$ telle que $T_i \to \infty$ quand $i \to \infty$ et pour chaque i, $(X_{T_i \wedge t})$ est une martingale.

Intégrales stochastiques par rapport aux mesures de Poisson Dans la section précedente nous avons déjà rencontré, dans la décomposition de Lévy-Itô, les intégrales des fonctions déterministes par rapport aux mesures de Poisson et mesures de Poisson compensées. Dans cette section nous allons généraliser cette intégrale aux fonctions stochastiques.

Soit M une mesure aléatoire de Poisson sur $[0,T] \times \mathbb{R}$ d'intensité μ. L'intensité μ est supposée σ-finie: il existe une suite $U_n \uparrow \mathbb{R}$ avec $\mu([0,t] \times U_n) < \infty$ pour tout t. Typiquement, M est la mesure de sauts d'un processus de Lévy. On voudrait définir l'intégrale de M ou de sa version compensée par rapport à une fonction prévisible $\phi : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$, c'est-à-dire, une fonction qui satisfait:

(i) Pour tout t, $(\omega, x) \mapsto \phi(\omega, t, x)$ est $\mathcal{F}_t \times \mathcal{B}(\mathbb{R})$-mesurable.

(ii) Pour tous (ω, x) $t \mapsto \phi(\omega, t, x)$ est continue à gauche.

On va définir l'intégrale stochastique de ϕ par rapport à M dans deux situations différentes:

Cas 1: ϕ satisfait

$$\int_0^T \int_{\mathbb{R}} |\phi(t,y)| M(dt \times dy) < \infty \quad \text{p.s.}$$

Dans ce cas, on définit l'intégrale stochastique de ϕ par rapport à $M = \sum \delta_{(T_i, y_i)}$ via la somme absolument convergente

$$\int_0^t \int_{\mathbb{R}} \phi(t,y) M(dt \times dy) := \sum_{i:T_i \leq t} \phi(T_i, y_i).$$

Cas 2: ϕ est de carrée integrable, c'est-à-dire, il satisfait

$$E \int_0^T \int_{\mathbb{R}} \phi^2(t,y) \mu(dt \times dy) < \infty$$

Dans ce cas, la construction est plus compliquée, et il faut utiliser la théorie L^2 et l'extension par continuité. On définit des fonctions prévisibles simples $\phi : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$ par

$$\phi(t,y) = \sum_{j=1}^m \phi_{0j} 1_{t=0} 1_{A_j}(y) + \sum_{i=1}^n \sum_{j=1}^m \phi_{ij} 1_{(T_i, T_{i+1}]}(t) 1_{A_j}(y),$$

où $T_0 = 0$, $(T_i)_{i\geq 1}$ est une suite de temps d'arrêt, pour tout j, $A_j \in \mathcal{B}(\mathbb{R})$ tel que $\mu([0,t] \times A_j) < \infty$ pour tout t, et pour chaque i et j, ϕ_{ij} est borné et $\mathcal{F}_{T_i}$-mesurable. L'intégrale stochastique d'une fonction prévisible simple par rapport à M est défini par

$$\int_0^t \int_{\mathbb{R}} \phi(t,y) M(dt \times dy) := \sum_{i:T_i \leq t} \phi(T_i, y_i) \equiv \sum_{i,j=1}^{n,m} \phi_{ij} M((T_i \wedge t, T_{i+1} \wedge t] \times A_j)$$

De façon similaire, on peut définir l'intégrale par rapport à la mesure de Poisson compensée $\tilde{M} = M - \mu$:

$$\int_0^t \int_{\mathbb{R}} \phi(t,y)\tilde{M}(dt \times dy) := \sum_{i,j=1}^{n,m} \phi_{ij} \left\{M((T_i \wedge t, T_{i+1} \wedge t] \times A_j) - \mu((T_i \wedge t, T_{i+1} \wedge t] \times A_j)\right\}$$

On peut démontrer que ce dernier processus est une martingale et satisfait la "relation d'isométrie":

$$E[X_T^2] = E \int_0^T \int_{\mathbb{R}} \phi^2(t,y)\mu(dt \times dy).$$

Cette relation d'isométrie permet d'éteindre la notion de l'intégrale stochastique par rapport à une mesure de Poisson compensée à toute fonction prévisible ϕ de carrée intégrable. Ensuite, en utilisant la procedure de localisation, on peut éteindre cette définition de l'intégrale stochastique à toute fonction ϕ, adaptée et continu à gauche en t, mesurable en y telle que le processus

$$A_t := \int_0^t \int_{\mathbb{R}} \phi^2(s,y)\mu(ds \times dy)$$

est localement intégrable.

La notion de l'intégrale stochastique par rapport à une mesure aléatoire de Poisson est plus générale que celle de l'intégrale par rapport à un processus de Poisson: si S est un processus de Lévy constant par morceaux, on a

$$\int_0^T \phi_t dS_t = \sum \phi_t \Delta S_t = \int_0^T \int_{\mathbb{R}} \phi_t y J_S(dt \times dy),$$

i.e., l'intégrale par rapport à un processus s'écrit comme l'intégrale par rapport à une mesure d'une fonction bien particulière.

L'intégrale stochastique par rapport à une mesure de Poisson nous permettra de définir une nouvelle classe de processus, qui généralise la classe de processus de Lévy tout en gardant une structure facilement compréhensible; plusieurs auteurs appellent cette classe processus de Lévy-Itô. Rappellons qu'un processus de Lévy satisfait (avec un petit changement de notation)

$$X_t = \mu t + \sigma W_t + \int_0^t \int_{|x|>1} xM(ds \times dx) + \int_0^t \int_{|x|\leq 1} x\tilde{M}(ds \times dx),$$

où M est une mesure de Poisson d'intensité $dt \times \nu$. Pour un processus de Lévy-Itô, on autorise des coefficients non-constants et même aléatoires:

$$X_t = \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{|x|>1} \gamma_s(x)M(ds \times dx) + \int_0^t \int_{|x|\leq 1} \gamma_s(x)\tilde{M}(ds \times dx), \tag{5.17}$$

où μ et σ sont des processus adaptés localement bornés et $\gamma_t(x)$ est une fonction aléatoire adaptée et continu à gauche en t, mesurable en x, telle que le processus:

$$\int_{|x|\leq 1} \gamma_t^2(x)\nu(dx)$$

est localement borné.

L'avantage de cette classe consiste en des nouvelles propriétés de stabilité: si (X_t) est un processus de Lévy-Itô, alors pour une fonction $f \in C^2$, $(f(X_t))$ est également un processus de Lévy-Itô; une propriété que les processus de Lévy ne possedent pas.

Lorsque $\int_{|x|>1} |\gamma_t(x)|\nu(dx)$ est aussi localement borné, le processus X peut être décomposé en une partie 'martingale' et une partie 'drift':

$$X_t = \int_0^t (\mu_s + \int_{|x|>1} \gamma_t(x)\nu(dx))ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{\mathbb{R}} \gamma_s(x)\tilde{M}(ds \times dx),$$

et dans le cas purement martingale souvent rencontré dans les applications,

$$X_t = \int_0^t \sigma_s dW_s + \int_0^t \int_{\mathbb{R}} \gamma_s(x) \tilde{M}(ds \times dx), \tag{5.18}$$

et on a la relation d'isométrie

$$E[X_T^2] = E\left[\int_0^T \sigma_t^2 dt\right] + E\left[\int_0^T \int_{\mathbb{R}} \gamma_t^2(x)\nu(dx)dt\right]. \tag{5.19}$$

Formule de changement de variable pour les processus de Lévy-Itô En l'absence de sauts, la formule de changement de variable (formule d'Itô) pour une fonction $f \in C^2$ prend la forme

$$f(X_T) = f(X_0) + \int_0^T f'(X_t)dX_t + \frac{1}{2}\int_0^T f''(X_t)\sigma_t^2 dt.$$

Quand le processus n'a qu'un nombre fini de sauts sur $[0,T]$, on peut écrire $X_t := X_t^c + \sum_{s\leq t} \Delta X_s$ et appliquer la même formule entre les sauts:

$$f(X_T) = f(X_0) + \int_0^T f'(X_t)dX_t^c + \frac{1}{2}\int_0^T f''(X_t)\sigma_t^2 dt + \sum_{t\leq T:\Delta X_t\neq 0} \{f(X_t) - f(X_{t-})\}.$$

Quand les sauts sont en nombre infini, la dernière somme peut diverger, mais on a toujours

$$f(X_T) = f(X_0) + \int_0^T f'(X_{t-})dX_t + \frac{1}{2}\int_0^T f''(X_t)\sigma_t^2 dt + \sum_{t\leq T:\Delta X_t\neq 0} \{f(X_t) - f(X_{t-}) - f'(X_{t-})\Delta X_t\}. \tag{5.20}$$

Pour faire apparaître la décomposition (5.17) et montrer que la classe de processus de Lévy-Itô est stable par transformations avec des fonctions C^2, on réécrit l'expression ci-dessus:

$$\begin{aligned} f(X_T) = f(X_0) &+ \int_0^T \{\mu_t f'(X_t) + \frac{1}{2}\sigma_t^2 f''(X_t) \\ &+ \int_{|x|\leq 1} (f(X_t + \gamma_t(x)) - f(X_t) - \gamma_t(x)f'(X_t))\nu(dx)\}dt \\ &+ \int_0^T f'(X_t)\sigma_t dW_t + \int_0^T \int_{|x|\leq 1} (f(X_{t-} + \gamma_t(x)) - f(X_{t-}))\tilde{M}(dt \times dx) \\ &+ \int_0^T \int_{|x|>1} (f(X_{t-} + \gamma_t(x)) - f(X_{t-}))M(dt \times dx) \end{aligned}$$

Exercice 17. Demontrer que pour un processus prévisible simple ϕ de la forme (5.15), et une martingale de carrée intégrable M, $\int_0^\cdot \phi_t dM_t$ est également une martingale et

$$E\left(\int_0^T \phi_t dM_t\right)^2 \leq \sup_{0\leq t\leq T, \omega\in\Omega} \phi_t^2 E[M_T^2].$$

Indication: utiliser le théorème d'arrêt de Doob. Soit (X_t) une martingale par rapport à la filtration $(\mathcal{F}_t)$ et soit S et T temps d'arrêt bornés avec $S \leq T$ p.s. Alors

$$E[X_T|\mathcal{F}_S] = X_S, \quad p.s.$$

Exercice 18. La variation quadratique ou "crochet droit" d'une sémimartingale peut être définie par

$$[X]_t := X_t^2 - X_0^2 - 2\int_0^t X_{s-} dX_s.$$

- Calculer la variation quadratique pour un processus de Lévy-Itô général, pour un processus de Lévy et pour le processus de Poisson.
- Montrer que si X est un processus de Lévy-Itô tel que $[X]_t \equiv t$ et X est une martingale alors X est un mouvement brownien standard.

Exercice 19. Soit X un processus de Lévy-Itô de la forme (5.17) avec coefficients μ, σ et γ déterministes et bornés. En appliquant la formule de changement de variable (5.20) à la fonction $f(x) = e^{iux}$, montrer que la fonction caractéristique de X_T peut être obtenue par une formule de Lévy-Khintchine généralisée.

Exercice 20. Soit X un processus de Lévy-Itô de la forme (5.17) tel que

$$\mu_t + \frac{\sigma_t^2}{2} + \int_{\mathbb{R}} (e^{\gamma_t(x)} - 1 - \gamma_t(x) 1_{|x|\leq 1})\nu(dx) = 0$$

p.s. pour tout t. En appliquant la formule de changement de variable (5.20), montrer que e^{X_t} s'écrit sous la forme (5.18) avec les coefficients à préciser.

En supposant que σ_t et $\int_{\mathbb{R}} (e^{\gamma_t(x)} - 1)^2 \nu(dx)$ sont bornés p.s. par une constante C, montrer, en utilisant la relation d'isométrie (5.19) et le lemme de Gronwall, que (e^{X_t}) est une martingale de carré intégrable.

Indication: utiliser le lemme de Gronwall. Soit ϕ une fonction positive localement bornée sur $\mathbb{R}^+$ telle que

$$\phi(t) \leq a + b\int_0^t \phi(s) ds$$

pour tout t et deux constantes a et $b \geq 0$. Alors $\phi(t) \leq ae^{bt}$.

5.5 Exponentielle stochastique d'un processus avec sauts

Proposition 11 (Exponentielle stochastique)**.** *Soit $(X)_{t\geq 0}$ un processus de Lévy-Itô de coefficient de volatilité σ. Il existe un unique processus càdlàg $(Z)_{t\geq 0}$ tel que*

$$dZ_t = Z_{t-} dX_t \qquad Z_0 = 1. \tag{5.21}$$

Z est explicitement donné par:

$$Z_t = e^{X_t - \frac{1}{2}\int_0^t \sigma_s^2 ds} \prod_{0\leq s\leq t} (1 + \Delta X_s) e^{-\Delta X_s}. \tag{5.22}$$

Preuve. Soit

$$V_t = \prod_{0\leq s\leq t; \Delta X_s \neq 0} (1 + \Delta X_s) e^{-\Delta X_s}.$$

La première étape est de montrer que ce processus existe et est de variation finie. On décompose V comme un produit de deux termes: $V_t = V_t' V_t''$, où

$$V_t' = \prod_{\substack{0\leq s\leq t \\ |\Delta X_s|\leq 1/2}} (1 + \Delta X_s) e^{-\Delta X_s} \quad \text{et} \quad V_t'' = \prod_{\substack{0\leq s\leq t \\ |\Delta X_s|> 1/2}} (1 + \Delta X_s) e^{-\Delta X_s}.$$

V'' pour tout t est un produit d'un nombre fini de facteurs, donc, il n'y a pas de problèmes de convergence. V' est positif et on peut prendre son logarithme.

$$\ln V_t' = \sum_{0\leq s\leq t; |\Delta X_s|\leq 1/2} (\ln(1 + \Delta X_s) - \Delta X_s).$$

Chaque terme de la somme satisfait

$$0 > \ln(1+\Delta X_s) - \Delta X_s > -\Delta X_s^2.$$

La série est donc décroissante et bornée inférieurement par $-\sum_{0\le s\le t}\Delta X_s^2$, ce qui est fini pour tout processus de Lévy-Itô. Donc, $(\ln V_t')$ existe et est un processus décroissant. Ceci implique que (V_t) existe et a des trajectoires de variation finie.

La deuxième étape est d'appliquer la formule d'Itô à la fonction $Z_t \equiv f(t, X_t, V_t) \equiv e^{X_t - \frac{1}{2}\int_0^t \sigma_s^2 ds} V_t$, ce qui donne

$$\begin{aligned} dZ_t &= -\frac{\sigma_s^2}{2} e^{X_{t-}-\frac{1}{2}\int_0^t \sigma_s^2 ds} V_{t-} dt + e^{X_{t-}-\frac{1}{2}\int_0^t \sigma_s^2 ds} V_{t-} dX_t + e^{X_{t-}-\frac{1}{2}\int_0^t \sigma_s^2 ds} dV_t \\ &+ \frac{\sigma_s^2}{2} e^{X_{t-}-\frac{1}{2}\int_0^t \sigma_s^2 ds} V_{t-} dt + e^{X_t-\frac{1}{2}\int_0^t \sigma_s^2 ds} V_t - e^{X_{t-}-\frac{1}{2}\int_0^t \sigma_s^2 ds} V_{t-} \\ &- e^{X_{t-}-\frac{1}{2}\int_0^t \sigma_s^2 ds} V_{t-}\Delta X_t - e^{X_{t-}-\frac{1}{2}\int_0^t \sigma_s^2 ds}\Delta V_t. \end{aligned}$$

Puisque V est un processus de saut pur,

$$dV_t \equiv \Delta V_t = V_{t-}(e^{-\Delta X_t}(1+\Delta X_t) - 1).$$

En substituant cela dans l'égalité précedente et en rassemblant les termes, on trouve (5.21).

Pour comprendre pourquoi la solution est unique, observons que si $(Z_t^{(1)})$ et $(Z_t^{(2)})$ verifient l'équation (5.21), alors leur différence $\tilde{Z}_t = Z_t^{(1)} - Z_t^{(2)}$ vérifie la même équation avec condition initiale $\tilde{Z}_0 = 0$. De la forme de cette équation il est clair que si la solution devient nulle à un point donné, elle reste nulle. □

Le processus Z est appellé exponentielle stochastique ou exponentielle de Doléans-Dade de X noté par $Z = \mathcal{E}(X)$.

Relation entre exponentielle ordinaire et exponentielle stochastique Il est clair que l'exponentielle ordinaire et l'exponentielle stochastique d'un processus sont deux processus stochastiques différents. Par exemple, contrairement à l'exponentielle ordinaire e^{X_t} qui est manifestement positive, l'exponentielle stochastique $Z = \mathcal{E}(X)$ n'est pas nécessairement positive. Il est facile de voir que l'exponentielle stochastique de X reste positive si tous les sauts de X sont plus grands que -1, ce qui revient à dire (pour un processus de Lévy) que la mesure de Lévy satisfait $\nu((-\infty,-1]) = 0$.

Il est donc naturel de demander, lequel de deux processus convient mieux pour construire des modèles financiers. Le résultat suivant, dû à Goll et Kallsen [36] montre que les deux approches sont équivalents: les deux opérations correspondent à la même classe de processus positifs.

Proposition 12 (Relation entre exponentielles ordinaires et stochastiques)**.**

1. *Soit $(X_t)_{t\ge 0}$ un processus de Lévy de triplet caractéristique (σ^2, ν, γ) et $Z = \mathcal{E}(X)$ son exponentielle stochastique. Si $Z > 0$ p.s. alors il existe un autre processus de Lévy $(L_t)_{t\ge 0}$ de triplet caractéristique $(\sigma_L^2, \nu_L, \gamma_L)$ tel que $Z_t = e^{L_t}$ où*

$$\begin{aligned} L_t &= \ln Z_t = X_t - \frac{\sigma^2 t}{2} + \sum_{0\le s\le t}\left\{\ln(1+\Delta X_s) - \Delta X_s\right\}. \qquad (5.23) \\ \sigma_L &= \sigma, \\ \nu_L(A) &= \nu(\{x : \ln(1+x) \in A\}) = \int 1_A(\ln(1+x))\nu(dx), \qquad (5.24) \\ \gamma_L &= \gamma - \frac{\sigma^2}{2} + \int \nu(dx)\left\{\ln(1+x)1_{[-1,1]}(\ln(1+x)) - x1_{[-1,1]}(x)\right\}. \end{aligned}$$

2. *Soit $(L_t)_{t\ge 0}$ un processus de Lévy de triplet caractéristique $(\sigma_L^2, \nu_L, \gamma_L)$ et $S_t = \exp L_t$ son exponentielle. Il existe un autre processus de Lévy $(X_t)_{t\ge 0}$ tel que S_t est l'exponentielle stochastique de X: $S = \mathcal{E}(X)$ où*

$$X_t = L_t + \frac{\sigma^2 t}{2} + \sum_{0\le s\le t}\left\{e^{\Delta L_s} - 1 - \Delta L_s\right\}. \qquad (5.25)$$

Le triplet (σ^2, ν, γ) de X est donné par:

$$\begin{aligned}
&\sigma = \sigma_L,\\
&\nu(A) = \nu_L(\{x : e^x - 1 \in A\}) = \int 1_A(e^x - 1)\nu_L(dx),\\
&\gamma = \gamma_L + \frac{\sigma_L^2}{2} + \int \nu_L(dx)\left\{(e^x - 1)1_{[-1,1]}(e^x - 1) - x1_{[-1,1]}(x)\right\}.
\end{aligned} \tag{5.26}$$

Preuve. 1. La condition $Z > 0$ p.s. est équivalente à $\Delta X_s > -1$ pour tout s p.s., ce qui permet de prendre le logarithme. Dans la preuve de proposition 11 nous avons vu que la somme $\sum_{0\le s\le t}\left\{\ln(1 + \Delta X_s) - \Delta X_s\right\}$ converge vers un processus de variation finie. Il est alors clair que L est un processus de Lévy et que $\sigma_L = \sigma$. De plus, $\Delta L_s = \ln(1 + \Delta X_s)$ pour tout s. Ceci implique

$$J_L([0,t] \times A) = \int_{[0,t]\times\mathbb{R}} 1_A(\ln(1 + x))J_X(ds\ dx)$$

et donc $\nu_L(A) = \int 1_A(\ln(1+x))\nu(dx)$. Il reste à calculer γ_L. En substituant la décomposition de Lévy-Itô pour (L_t) et (X_t) dans (5.23), on obtient

$$\begin{aligned}
\gamma_L t - \gamma t + \frac{\sigma^2 t}{2} \quad &+ \quad \int_{s\in[0,t],|x|\le 1} x\tilde{J}_L(ds\ dx) + \int_{s\in[0,t],|x|>1} xJ_L(ds\ dx)\\
&- \quad \int_{s\in[0,t],|x|\le 1} x\tilde{J}_X(ds\ dx) - \int_{s\in[0,t],|x|>1} xJ_X(ds\ dx)\\
&- \quad \sum_{0\le s\le t}\left\{\ln(1 + \Delta X_s) - \Delta X_s\right\} = 0.
\end{aligned}$$

En observant que

$$\begin{aligned}
&\int_{s\in[0,t],|x|\le 1} x(J_X(ds\ dx) - J_L(ds\ dx))\\
&\qquad = \sum_{0\le s\le t}\left(\Delta X_s 1_{[-1,1]}(\Delta X_s) - \ln(1 + \Delta X_s)1_{[-1,1]}(\ln(1 + \Delta X_s))\right)
\end{aligned}$$

converge, on peut séparer l'expression ci-dessus en une partie 'sauts' et une partie 'drift', donc chacune doit être égale à zéro. Pour la partie drift on obtient:

$$\gamma_L - \gamma + \frac{\sigma^2}{2} - \int_{-1}^{1}\{x\nu_L(dx) - x\nu(dx)\} = 0,$$

ce qui donne la formule pour γ_L après un changement de variable.

2. Les sauts de S_t sont donnés par $\Delta S_t = S_{t-}(\exp(\Delta L_t)-1)$. Si X est un processus de Lévy tel que $S = \mathcal{E}(X)$ alors puisque $dS_t = S_{t-}dX_t$ on a $\Delta S_t = S_{t-}\Delta X_t$ et donc $\Delta X_t = \exp(\Delta L_t) - 1$ et ν est donné par (5.26). En particulier $\Delta X_t > -1$ p.s. et il est facile à vérifier que $\ln \mathcal{E}(X)$ est un processus de Lévy avec les mêmes caractéristiques que L seulement si les caractéristiques de X sont données par (5.26). Inversement, si X est un processus de Lévy avec caractéristiques données par (5.26), (5.22) permet de vérifier que $\mathcal{E}(X) = \exp L_t$. □

En vu de ce résultat, et du fait que les formules qui font intervenir l'exponentielle ordinaire sont plus lisibles, cette dernière est plus souvent utilisée pour construire les modèles de prix. Cependant, dans certaines situations comme le montre l'exemple suivant, l'exponentielle stochastique est mieux adaptée.

Stratégie CPPI avec sauts La stratégie CPPI est une stratégie d'*assurance de portefeuille* permettant (en théorie) de maintenir la valeur du portefeuille au-dessus d'un seuil donné tout en profitant d'une evolution favorable de la bourse.

Pour fixer les idées, nous supposerons que l'investisseur souhaite garantir le capital N à l'échéance T (N peut être plus grand ou plus petit que l'investissement initial). Pour cela il doit maintenir la valeur du portefeuille V_t au-dessus du *plancher* B_t à chaque date, où B_t est égal au prix du zéro-coupon de nominal N et d'échéance T. La difference $C_t = V_t - B_t$ est appellée le *coussin*. La stratégie CPPI consiste alors à utiliser l'algorithme suivant:

- A chaque date t, si $V_t > B_t$, investir mC_t en l'actif risqué, où $m > 1$ est le *multiplicateur* de la stratégie, et le reste en zéro-coupons d'échéance T.
- Si $V_t \leq B_t$, investir tout en zéro-coupons d'échéance T.

Si le prix de l'actif risqué S est un processus aux trajectoires continues, le portefeuille reste toujours au-dessus du plancher, et l'évolution du coussin est donnée par

$$\frac{dC_t}{C_t} = m\frac{dS_t}{S_t} + (1-m)rdt,$$

où r est le taux d'intérêt. Dans le modèle de Black-Scholes,

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t,$$

cette équation se résout explicitement et on obtient la valeur finale du portefeuille

$$V_T = N + (V_0 - Ne^{-rT}) \exp\left(rT + m(\mu - r)T + m\sigma W_T - \frac{m^2\sigma^2 T}{2}\right).$$

L'espérance du gain est donc

$$E[V_T] = N + (V_0 - Ne^{-rT}) \exp\left(rT + m(\mu - r)T\right).$$

Si $\mu > r$, on arrive à un paradoxe apparent: il n'y a pas de risque, et l'espérance du gain peut être rendu aussi grande que l'on souhaite en choisissant un multiplicateur suffisamment élevé. Ce paradoxe se résout facilement en introduisant les sauts dans les trajectoires du prix car dans ce cas, en augmentant le multiplicateur on augmente aussi la probabilité de perte.

Soit

$$\frac{dS_t}{S_{t-}} = rdt + dZ_t,$$

où Z_t est un processus de Lévy, et soit $\tau = \inf\{t : V_t \leq B_t\}$ la première date où le portefeuille passe au-dessous du coussin (on peut avoir $\tau = \infty$). Alors avant la date τ, le coussin satisfait

$$\frac{dC_t}{C_{t-}} = mdZ_t + rdt,$$

et le coussin actualisé $C_t^* := \frac{C_t}{e^{rt}}$ est donc donné par

$$C_t^* = \mathcal{E}(mZ)_t, \quad t < \tau.$$

Après la date τ, tout le portefeuille est investi en zéro coupons et le coussin actualisé reste constant. On a donc

$$C_t^* = \mathcal{E}(mZ)_{t\wedge\tau}.$$

La perte a lieu si, à un instant $t \leq T$, $C_t^* \leq 0$, ce qui peut arriver si et seulement si Z a un saut dans $[0,T]$ dont la taille est inférieure à $-1/m$. On obtient donc (voir exercice 11)

$$P[\exists t \in [0,T] : V_t \leq B_t] = 1 - \exp\left(-T\int_{-\infty}^{-1/m} \nu(dx)\right).$$

Exercice 21. Montrer que si X et Y sont deux processus de Lévy et Y est de saut pur alors

$$\mathcal{E}(X)\mathcal{E}(Y) = \mathcal{E}(X + Y + \sum \Delta X \Delta Y)$$

Exercice 22 (Propriété de martingale de l'exponentielle stochastique)**.** Soit X un processus de Lévy et une martingale. Montrer que $\mathcal{E}(X)$ est également une martingale.

Indication: Décomposer X en une somme d'un processus de Poisson composé et d'un processus de Lévy X' tel que $|\Delta X'| < 1$. Utiliser proposition 12 et l'exercice précedent.

5.6 Couverture dans les modèles avec sauts

Introduction Le problème de couverture se pose typiquement du point de vue du *vendeur d'options* (banque) qui voudrait minimiser son risque à maturité; pour cela il essaie de repliquer le pay-off Y de l'option avec un portefeuille autofinançant $V_T = V_0 + \int_0^T \phi_t dS_t$. Le prix de l'option est alors calculé comme le coût de sa couverture plus une provision pour le risque residuel (s'il y en a) plus la marge commerciale. Dans le modèle Black-Scholes, qui correspond à un marché complet, une replication parfaite est obtenu avec la stratégie $\phi_t = \frac{\partial C}{\partial S}(t, S_t)$ (couverture en delta).

Dans un marché incomplet, la replication exacte n'est pas possible et le problème de couverture devient un problème d'approximation du pay-off de l'option par le portefeuille de couverture. Pour faire face à ce problème, les praticiens calculent et neutralisent les sensibilités du portefeuille par rapport aux facteurs de risque additionnels (e.g. risque de volatilité ou risque de grands sauts).

$$\text{Delta} = \frac{\partial C(t, X_t)}{\partial x}: \quad \text{petites variations; couverture avec des actions}$$
$$\text{Gamma} = \frac{\partial^2 C(t, X_t)}{\partial x^2}: \quad \text{grandes variations; couverture avec des options liquides}$$

De manière plus systématique on peut essayer d'optimiser la stratégie de couverture en contrôlant l'erreur residuelle.

- La couverture par maximisation d'utilité consiste à chercher la stratégie de couverture qui maximise l'utilité terminale du vendeur de l'option

$$\max_\phi E\left[U(c + \int_0^T \phi_t dX_t - Y)\right] \quad (U \text{ croissante concave}).$$

  Un inconvenient de cette approche est qu'elle correspond à une règle de pricing et couverture non-linéaire: la couverture pour un portefeuille contenant une option A et une option B ne coïncide pas avec la couverture de A plus la couverture de B.

- La couverture quadratique donne, quant à elle, un ratio de couverture linéaire. Elle consiste à minimiser la distance L^2 entre le pay-off et la valeur terminale du portefeuille de couverture:

$$\min_\phi E\left(c + \int_0^T \phi_t dX_t - Y\right)^2$$

  Le portefeuille de couverture optimal (s'il existe) est la projection L^2 de Y sur le sous-espace (lineaire) d'actifs replicables. Par contre, elle penalise les gains et les pertes de la même façon.

Dans la suite de cette section on va se concentrer sur la couverture quadratique. De plus, on supposera que les prix de tous les actifs sont des martingales. Ceci simplifie considerablement le calcul et donne une bonne approximation: la prise en compte du drift du processus ne modifie pas la stratégie de couverture dans un marché complet et mène a une correction de 2ème ordre dans un marché incomplet.

Le modèle On supposera que les prix de m actifs cotés $X_t \in \mathbb{R}^m$ sont des processus de Lévy-Itô martingales

$$X_t = X_0 + \int_0^t \sigma_s dW_s + \int_0^t \int_{\mathbb{R}} \gamma_s(z) \tilde{J}(ds \times dz),$$

où

- W est un mouvement Brownien d-dimensionnel.

- J est une mesure aléatoire de Poisson d'intensité $dt \times \nu$.
- σ et γ: sont des processus càdlàg adaptés qui remplissent les conditions d'intégrabilité:

$$\|\gamma_s(z)\|^2 \leq \rho(z)A_s \quad \text{with} \quad \int \rho(z)\nu(dz) < \infty$$
$$E\int_0^T (\|\sigma_s\|^2 + A_s)ds < \infty$$

Le pay-off de l'option Y est de carré intégrable: $E[Y^2] < \infty$ et on supposera que la martingale

$$Y_t = E[Y|\mathcal{F}_t]$$

admet une representation de type Lévy-Itô.

$$Y_t = Y_0 + \int_0^t \sigma_s^0 dW_t + \int_0^t \int_{\mathbb{R}} \gamma_s^0(z)\tilde{J}(ds \times dz)$$

Exemple 6. Le sous-jacent suit un modèle exp-Lévy $X_t = X_0 e^{Z_t}$, ce qui implique

$$X_t = X_0 + \int_0^t \sigma X_s dW_s + \int_{[0,t]\times\mathbb{R}} X_{s-}(e^z - 1)\tilde{J}(ds \times dz),$$

et l'actif à couvrir est une option européenne

$$C(t,x) = E[H(X_T^{(t,x)})]$$
$$dC(t,X_t) = \frac{\partial C(t,X_t)}{\partial X}\sigma X_t dW_t + \int_{\mathbb{R}} (C(t,X_{t-}e^z) - C(t,X_{t-}))d\tilde{J}(ds \times dz).$$

Le ratio optimal de couverture L'erreur de couverture est définie par

$$\epsilon_T(c,\phi) = c + \int_0^T \phi_t dX_t - Y$$
$$= c - E[Y] + \int_0^T (\phi_t\sigma_t - \sigma_t^0)dW_t + \int_{[0,T]\times\mathbb{R}} (\phi_t\gamma_t(z) - \gamma_t^0(z))\tilde{J}(dt \times dz).$$

Une stratégie admissible est une stratégie ϕ telle que $\int_0^{\cdot} \phi_t dX_t$ est une martingale de carrée intégrable. Le ratio optimal de couverture $(\hat{c}, \hat{\phi})$ est la solution de

$$E[\epsilon_T(\hat{c},\hat{\phi})^2] = \inf_{(c,\phi)} E[(\epsilon_T(c,\phi))^2].$$

Pour résoudre ce problème d'optimisation, on commence par remarquer que $\hat{c}$ doit être égal à $E[Y]$ car

$$E[(\epsilon_T(c,\phi))^2] = (c - E[Y])^2 + E\left(E[Y] - Y + \int_0^T \phi_t dX_t\right)^2.$$

Ensuite, on utilise la relation d'isométrie pour les processus de Lévy-Ito.

$$E[\epsilon(\hat{c},\phi)_T^2] = \int_0^T dt E\|\phi_t\sigma_t - \sigma_t^0\|^2 + \int_0^T dt \int \nu(dz)E(\phi_t\gamma_t(z) - \gamma_t^0(z))^2$$

ce qui est minimisé par

$$\hat{\phi}_t = M_t^{-1}\left(\sigma_t^0\sigma_t^* + \int_{\mathbb{R}} \nu(dz)\gamma_t^0(z)\gamma_t(z)^*\right),$$
$$\text{si} \quad M_t = \sigma_t\sigma_t^* + \int_{\mathbb{R}} \nu(dz)\gamma_t(z)\gamma_t(z)^* \quad \text{est non-singulière.}$$

Couverture avec des actions dans un modèle exp-Lévy

$$\begin{aligned} dX_t &= X_t \sigma dW_t + \int_{\mathbb{R}} X_{t-}(e^z - 1) d\tilde{J}(dt \times dz), \\ Y_t &= C(t, X_t) \end{aligned}$$

Le ratio optimal de couverture est donné par

$$\phi_t^{opt} = \frac{\sigma^2 \frac{\partial C}{\partial X}(t, X_t) + \frac{1}{X_t} \int \nu(dz)(e^z - 1)[C(t, X_t e^z) - C(t, X_t)]}{\sigma^2 + \int (e^z - 1)^2 \nu(dz)}$$

L'erreur de couverture s'annule si pour presque tout t, $\exists k$:

$$(\sigma X_t \frac{\partial C}{\partial X}, (C(t, X_t e^z) - C(t, X_t))_{z \in \operatorname{supp} \nu}) = k(\sigma X_t, (X_t(e^z - 1))_{z \in \operatorname{supp} \nu})$$

Ceci se produit dans les deux cas suivants:

- Mouvement brownien: $\nu = 0$.
- Processus de Poisson: $\sigma = 0, \nu = \delta_{x_0}(x)$.

Dans tous les autres cas, le marché est incomplet.

Couverture en delta vs. couverture optimale Quelle est la distance entre la couverture en delta et la couverture optimale? Si le processus sous-jacent a une composante de diffusion et les sauts sont petits, alors

$$\phi_t = \frac{\partial C}{\partial X} + \frac{X_t}{2\Sigma^2} \frac{\partial^2 C}{\partial X^2} \int \nu(dz)(e^z - 1)^3.$$

où

$$\Sigma^2 = \sigma^2 + \int (e^z - 1)^2 \nu(dz).$$

La couverture optimale apparait alors comme une petite correction (voir fig. 5.1), qui est souvent négative sur les marchés d'options car les sauts sont négatifs. On réduit le ratio de couverture en anticipant les sauts négatifs dans les prix d'actions. Pour les processus de saut pur tel que le modèle variance gamma, la correction peut être grande car $\frac{\partial^2 C}{\partial X^2}$ peut ne pas être défini (voir fig. 5.2).

En terme d'erreur de couverture, si les sauts sont petits, la couverture en delta marche bien et sa performance est proche de l'optimale. Cette situation est illustrée sur le graphique de gauche de fig. 5.3. En revanche, si la composante de sauts est forte, la stratégie optimale est clairement supérieure à la couverture en delta, mais l'erreur residuelle est assez importante pour les deux stratégies (fig. 5.3, graphique de droite). Dans ce dernier cas, la performace peut être améliorée en ajoutant des options liquides dans le portefeuille de couverture (fig. 5.4).

Exercice 23 (Couverture dans un modèle basé sur l'exponentielle stochastique)**.** Calculer le ratio optimal de couverture (avec actions) dans un modèle où le sous-jacent est donné par l'exponentielle stochastique d'un processus de Lévy: $S_t = S_0 \mathcal{E}(X)_t$.

Exercice 24 (Ratios de couverture dans le modèle diffusion + Poisson)**.** On suppose que le sous-jacent est décrit par

$$S_t = S_0 \exp(\mu t + \sigma W_t - \delta N_t),$$

où N est un processus de Poisson composé d'intensité λ.

- Calculer le ratio optimal pour la couverture d'une option européenne (dont le prix sera noté par C) avec des actions et une autre option européenne (dont le prix sera noté par C_0).
- Montrer que dans ce cas la replication est exacte.

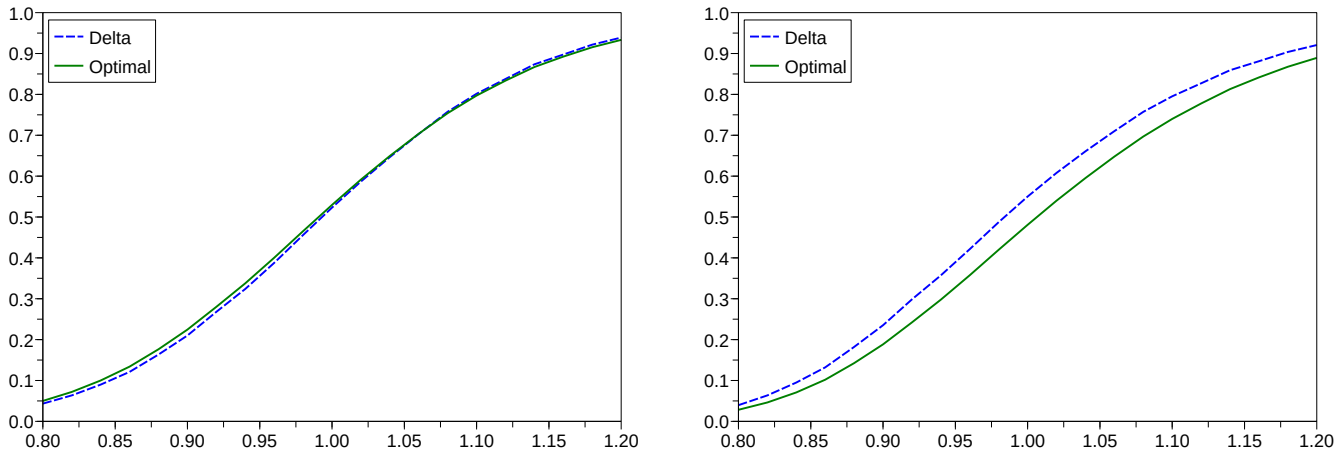


Figure 5.1: La différence entre la couverture en delta et la couverture optimale dans le modèle de Kou (diffusion plus sauts positif et négatifs distribués selon la loi exponentielle). Gauche : paramètres estimés sur les données de marché. Droite: tous les sauts sont négatifs.

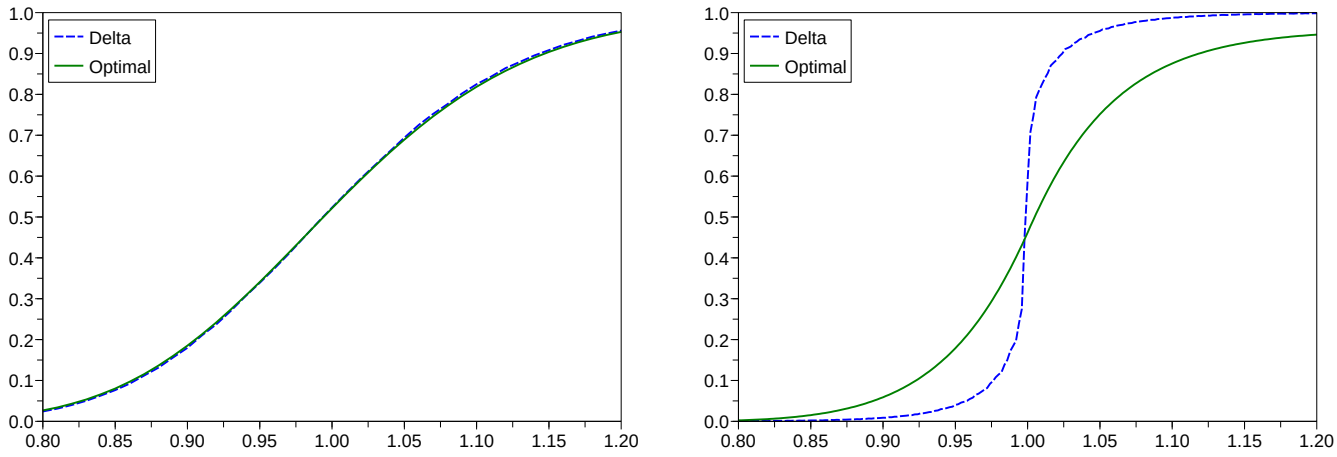


Figure 5.2: La différence entre la couverture en delta et la couverture optimale dans le modèle variance gamma. Gauche : paramètres estimés sur les données de marché, échéance $T = 1$ mois. Droite la taille moyenne de saut est de 7%, échéance $T = 2$ jours.

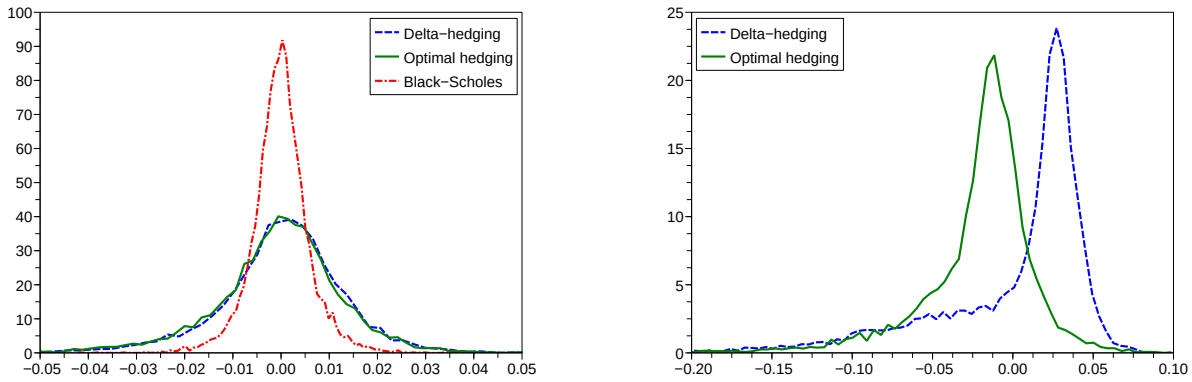


Figure 5.3: Histogrammes de l'erreur pour la couverture d'un put Européen hors-la-monnaie ($K = 90\%$, $T = 1$ an). Gauche: modèle de Kou, paramètres estimés sur les données Microsoft. Droite: modèle de Kou, grands sauts négatifs peu fréquents (taille moyenne de sauts 10%.)

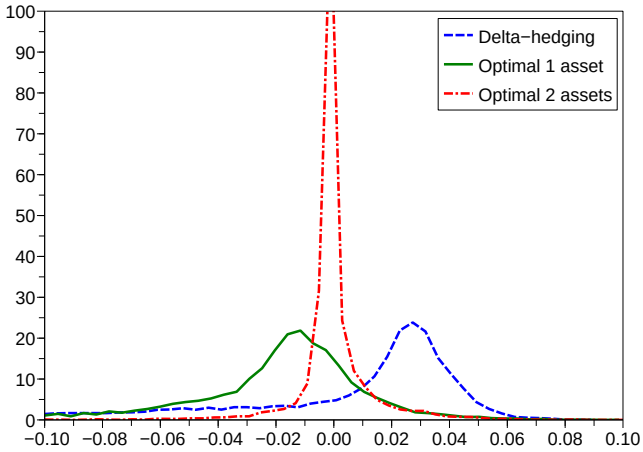


Figure 5.4: Pour reduire l'erreur de couverture en présence de grands sauts, des options liquides peuvent être utilisées pour la couverture. Le graphique montre l'histogramme de l'erreur pour la couverture d'un put européen hors la monnaie (strike $K = 90\%$) avec des actions et un call européen (strike $K = 110\%$).

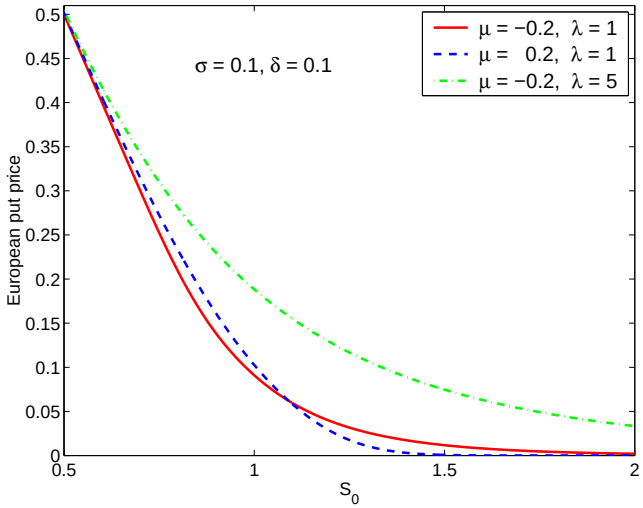


Figure 5.5: Le prix du put européen dans le modèle de Merton pour trois choix de paramètres. Les autres paramètres sont $K = 1$, $T = 1$, $r = 0$.

- Montrer que lorsque $\delta \to 0$, on retrouve les formules de couverture delta-gamma

$$\gamma_t = \frac{\partial^2 C}{\partial S^2} \Big/ \frac{\partial^2 C_0}{\partial S^2} \quad \text{(nombre d'options } C_0\text{)}$$
$$\Delta_t = \frac{\partial C}{\partial S} - \gamma_t \frac{\partial C_0}{\partial S} \quad \text{(nombre d'actions)}$$

5.7 Valorisation d'options dans les modèles exp-Lévy

Méthodes de transformée de Fourier pour les options européennes Dans les modèles exp-Lévy, les prix d'options européennes peuvent être calculés avec la méthode de transformée de Fourier décrite dans la section suivante. Cette méthode s'applique à tous les modèles où la fonction caractéristique du log-prix est connue ou facile à calculer. C'est le cas par exemple pour les modèles exp-Lévy mais aussi pour une classe plus général de processus *affines* [25, 26], qui contient en particulier le modèle de Bates dont on parlera à la fin de ce chapitre. Le graphique 5.5 trace les prix d'une option put en fonction du strike dans le modèle de Merton avec paramètres différents.

Equations intégro-différentielles pour les options exotiques Pour certaines options dont le pay-off dépend de la trajectoire empruntée par le processus de prix, telles que les options à barrière, le prix peut être exprimé comme solution d'une équation similaire à l'équation de Black-Scholes

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} = rC - rS\frac{\partial C}{\partial S}.$$

Soit X un processus de Lévy tel que e^X est une martingale sous la probabilité Q (probabilité risque-neutre) et le prix de sous-jacent est donné par $S_t = S_0 e^{rt+X_t}$. Alors, le prix à l'instant t d'une option européenne

$$P_t = e^{-r(T-t)} E[(S_T - K)^+ | \mathcal{F}_t]$$

peut être exprimée comme fonction déterministe de t et S_t: $P_t = P(t, S_t)$ et de plus $e^{-rt}P(t, S_t)$ est une martingale.

De manière similaire, pour une option up-and-out on a

$$P_t^B = e^{-r(T-t)} E[(S_T - K)^+ 1_{\max_{0\le s\le T} S_s < B} | \mathcal{F}_t].$$

Le processus $e^{-rt}P_t^B$ est donc une martingale. Soit maintenant $P^B(t,S)$ une fonction déterministe définie par

$$P^B(t,S) = e^{-r(T-t)}E[g(S_{T\wedge\tau_t})|S_t = S],$$

où $g(S) = (S-K)^+ 1_{S<B}$ et $\tau_t = \inf\{s \geq t : X_s \geq B\}$. Alors, $P_t^B = P^B(t,S_t)$ si la barrière n'a pas encore été atteinte.

Sous l'hypothèse que e^X est une martingale, la dynamique de S s'écrit

$$dS_t = rS_t dt + S_t \sigma dW_t + \int_{\mathbb{R}} S_{t-}(e^x - 1)\tilde{J}_X(dt \times dx),$$

où J_X est la mesure de sauts compensée de X et σ sa volatilité. Prenons alors une fonction suffisamment régulière $\tilde{P}(t,S_t)$. En appliquant la formule d'Itô à cette fonction, on trouve alors

$$\begin{aligned} d\tilde{P}(t,S_t) = \Bigg\{ &\frac{\partial \tilde{P}}{\partial t} + rS\frac{\partial \tilde{P}}{\partial S} + \frac{1}{2}S_t^2\sigma^2\frac{\partial^2 \tilde{P}}{\partial S^2} \\ &+ \int_{\mathbb{R}} \left(\tilde{P}(t,S_t e^x) - \tilde{P}(t,S_t) - S_t(e^x-1)\frac{\partial \tilde{P}}{\partial S} \right) \nu(dx) \Bigg\} \\ &+ \frac{\partial \tilde{P}}{\partial S}\sigma S_t dW_t + \int_{\mathbb{R}} (\tilde{P}(t,S_t e^x) - \tilde{P}(t,S_t))\tilde{J}_X(dt \times dx). \end{aligned}$$

Nous pouvons alors énoncer la proposition suivante (les résultats de ce genre sont souvent appellés théorèmes de verification).

Proposition 13. *Soit $P(t,S)$ une fonction 1 fois dérivable en t et 2 fois dérivable en S avec la dérivée $\frac{\partial P}{\partial S}$ bornée, qui satisfait l'équation*

$$\frac{\partial P}{\partial t} + \frac{1}{2}S^2\sigma^2\frac{\partial^2 P}{\partial S^2} + \int_{\mathbb{R}} \left(P(t,Se^x) - P(t,S) - S(e^x-1)\frac{\partial P}{\partial S} \right) \nu(dx) = rP - rS\frac{\partial P}{\partial S}$$

avec la condition terminale $P(T,S) = (S-K)^+$. Alors le prix à l'instant t d'une option call européenne d'échéance T et de strike K est donné par $P(t,S_t)$.

Le prix d'une option européenne dans un modèle avec sauts résout alors une équation similaire à l'EDP Black-Scholes, mais qui contient un terme de correction dépendant de la mesure de sauts. Les équations de ce type s'appellent équations intégro-différentielles. Il reste alors à demontrer que l'équation ci-dessus admet des solutions regulières, et pour les options européennes c'est un exercice relativement facile [16]. On pourrait formellement demontrer que le prix d'une option up-and-out avec la barrière B résout la même équation avec la condition supplémentaire $P^B(t,S) = 0$ pour $S \geq B$, mais dans ce cas le problème de régularité est vraiment difficile, et on est obligé de recourir aux solutions faibles au sens de viscosité [16].

Exercice 25 (Options forward start)**.** Les options forward start sont les options dont le strike est déterminé à une date future selon une règle spécifique, par exemple,

$$H_T = (S_T - mS_{T_0})^+, \tag{1}$$

où $T_0 < T$ est une date future et m est un nombre fixé dans le contrat (moneyness de l'option). On se place dans un modèle exponentielle Lévy sous la probabilité risque-neutre:

$$S_t = S_0 e^{rt + X_t},$$

où (X_t) est un processus de Lévy, et on suppose que les prix des options Européennes de tous les strikes et toutes les maturités sont observables sur le marché.

1. Calculer (exprimer à partir des prix des européens) le prix à l'instant $t = 0$ d'une option forward start avec pay-off donné par (1).

2. Une option cliquet (séquence d'options forward start) a un pay-off de

$$\sum_{i=1}^{12} \max(0, r_i), \quad r_i = S_{\frac{i}{12}} - S_{\frac{i-1}{12}}.$$

Calculer le prix à l'instant $t = 0$ de cette option. Est-ce que le résultat vous paraît réaliste? Conclure sur la pertinence des modèles exponentielle-Lévy pour le pricing des options cliquets.

5.8 Méthode de transformée de Fourier pour la valorisation d'options européennes

Let $\{X_t\}_{t\geq 0}$ be a stochastic process on $(\Omega, \mathcal{F}, P)$ such that e^{X_t} is a martingale. In order to compute the price of a call option

$$C(K) = e^{-rT} E[(S_0 e^{rT+X_T} - K)^+] = S_0 E\left[\left(e^{X_T} - \frac{Ke^{-rT}}{S_0}\right)^+\right] := S_0 E\left[\left(e^{X_T} - e^k\right)^+\right],$$

we would like to express its Fourier transform in adjusted log strike $k := \log \frac{Ke^{-rT}}{S_0}$ in terms of the characteristic function $\Phi_T(v)$ of X_T and then find the prices for a range of strikes by Fourier inversion. However we cannot do this directly because C as function of k is not integrable (it tends to a positive constant as $k \to -\infty$). The key idea is to instead compute the Fourier transform of the time value of the option (which we also renormalize to get rid of S_0), that is, the function

$$z_T(k) = E[(e^{X_T} - e^k)^+] - (1 - e^k)^+. \tag{5.27}$$

Proposition 14. *Let $\{X_t\}_{t\geq 0}$ be a stochastic process on $(\Omega, \mathcal{F}, P)$ such that e^{X_t} is a martingale and*

$$E[e^{(1+\alpha)X_t}] < \infty \quad \forall t \tag{5.28}$$

for some $\alpha > 0$. Then the Fourier transform of the time value of a call option is given by:

$$\zeta_T(v) := \int_{-\infty}^{+\infty} e^{ivk} z_T(k) dk = \frac{\Phi_T(v-i) - 1}{iv(1+iv)} \tag{5.29}$$

The option prices can be found by inverting the Fourier transform:

$$C(K) = (S_0 - Ke^{-rT})^+ + \frac{S_0}{2\pi} \int_{-\infty}^{+\infty} e^{-iv \log(Ke^{-rT}/S_0)} \frac{\Phi_T(v-i) - 1}{iv(1+iv)}.$$

Remarque 4. Since typically $\Phi_T(z) \to 0$ as $\Re z \to \infty$, $\zeta_T(v)$ will behave like $|v|^{-2}$ at infinity which means that the truncation error in the numerical evaluation of the inverse Fourier transform will be large. The reason of such a slow convergence is that the time value (5.27) is not smooth; therefore its Fourier transform does not decay sufficiently fast at infinity. For most models the convergence can be improved by replacing the time value with a smooth function of strike. Namely, instead of subtracting the intrinsic value of the option (which is non-differentiable) from its price, we suggest to subtract the Black-Scholes call price with a non-zero volatility (which is a smooth function). The resulting function will be both integrable and smooth. Suppose that the hypothesis of the above proposition is satisfied and denote

$$\tilde{z}_T(k) = E[(e^{X_T} - e^k)^+] - \frac{1}{S_0} c_{BS}^{\Sigma}(k),$$

where $c_{BS}^{\Sigma}(k)$ is the Black-Scholes price of a call option with volatility Σ and log-strike k for the same underlying value and the same interest rate. The above proposition implies that the Fourier transform of $\tilde{z}_T(k)$, denoted by $\tilde{\zeta}_T(v)$, satisfies

$$\tilde{\zeta}_T(v) = \frac{\Phi_T(v-i) - \Phi_T^{\Sigma}(v-i)}{iv(1+iv)}, \tag{5.30}$$

where $\Phi_T^\Sigma(v) = \exp(-\frac{\Sigma^2 T}{2}(v^2+iv))$. Since for most models found in the literature (except variance gamma) the characteristic function decays faster than every power of its argument at infinity, this means that the expression (5.30) will also decay faster than every power of v as $\Re v \to \infty$, and the integral in the inverse Fourier transform will converge very fast for *every* $\Sigma > 0$. The option prices can be found via

$$C(K) = \mathrm{Call}_{BS}^\Sigma(K) + \frac{S_0}{2\pi}\int_{-\infty}^{+\infty} e^{-iv\log(Ke^{-rT}/S_0)}\frac{\Phi_T(v-i)-\Phi_T^\Sigma(v-i)}{iv(1+iv)}.$$

Proof of Proposition 14. First, we note that since the discounted price process is a martingale, we can write

$$z_T(k) = \int_{-\infty}^{\infty}\mu_T(dx)(e^x-e^k)(1_{k\le x}-1_{k\le 0}),$$

where μ_T is the probability distribution of X_T. Condition (5.28) enables us to compute $\zeta_T(v)$ by interchanging integrals:

$$\begin{aligned}\zeta_T(v) &= \int_{-\infty}^{\infty}dk\int_{-\infty}^{\infty}\mu_T(dx)e^{ivk}(e^x-e^k)(1_{k\le x}-1_{k\le 0})\\ &= \int_{-\infty}^{\infty}\mu_T(dx)\int_x^0 e^{ivk}(e^k-e^x)dk\\ &= \int_{-\infty}^{\infty}\mu_T(dx)\left\{\frac{(1-e^x)}{iv+1}-\frac{e^x}{iv(iv+1)}+\frac{e^{(iv+1)x}}{iv(iv+1)}\right\}\end{aligned}$$

The first term in braces disappears due to the martingale condition and the other two, after computing the integrals, yield (5.29). □

Numerical Fourier inversion. Option prices can be computed by evaluating numerically the inverse Fourier transform of $\tilde\zeta_T$:

$$\tilde z_T(k) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-ivk}\tilde\zeta_T(v)dv \tag{5.31}$$

This integral can be efficiently computed using the Fast Fourier transform, an algoritm due to Cooley and Tukey [17] which allows to compute $F_0,\dots,F_{N-1}$, given by,

$$F_n = \sum_{k=0}^{N-1} f_k e^{-2\pi ink/N},\quad n=0\dots N-1,$$

using only $O(N\ln N)$ operations.

To approximate option prices, we truncate and discretize the integral (5.31) as follows:

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-ivk}\tilde\zeta_T(v)dv = \frac{1}{2\pi}\int_{-A/2}^{A/2}e^{-ivk}\tilde\zeta_T(v)dv+\varepsilon_T = \frac{A}{2\pi N}\sum_{m=0}^{N-1}w_m\tilde\zeta_T(v_m)e^{-ikv_m}+\varepsilon_T+\varepsilon_D, \tag{5.32}$$

where ε_T is the truncation error, ε_D is the discretization error, $v_m = -A/2+m\Delta$, $\Delta = A/(N-1)$ is the discretization step and w_m are weights, corresponding to the chosen integration rule (for instance, for the trapezoidal rule $w_0 = w_{N-1} = 1/2$ and all other weights are equal to 1). Now, setting $k_n = k_0 + \frac{2\pi n}{N\Delta}$ we see that the sum in the last term becomes a discrete Fourier transform:

$$\frac{A}{2\pi N}e^{ik_nA/2}\sum_{m=0}^{N-1}w_m e^{-ik_0v_m}f(k_m)e^{-2\pi inm/N}$$

The FFT algorithm allows to compute $\tilde z_T$ and therefore option prices for the log strikes $k_n = k_0+\frac{2\pi n}{N\Delta}$. The log strikes are thus equidistant with the step d satisfying

$$d\Delta = \frac{2\pi}{N}.$$

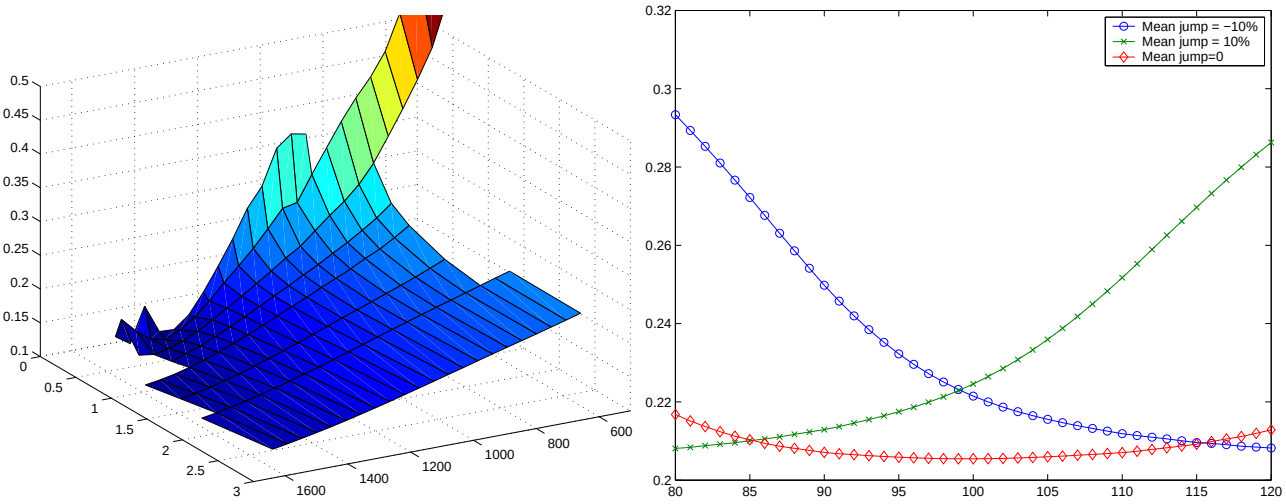


Figure 5.6: Left: implied volatilities of options on S&P 500 index as a function of their strikes and maturities. Right: implied volatilities as a function of strike for different values of the mean jump size in Merton jump diffusion model. Other parameters: volatility $\sigma = 0.2$, jump intensity $\lambda = 1$, jump standard deviation $\delta = 0.05$, option maturity $T = 1$ month.

Typically, k_0 is chosen so that the grid is centered around the money, Δ is fixed to keep the discretization error low, and N is adjusted to keep the truncation error low and have a sufficiently small step between strikes (increasing N reduces the truncation error and the distance between consecutive strikes at the same time). The option prices for the strikes not on the grid must be computed by interpolation. It should be noted that the FFT method should only be used when option prices for many strikes must be computed simultaneously (such as for calibration). If only a small number of strikes is needed, integration methods with variable step size usually have a better performance.

5.9 Calibration de modèles exp-Lévy

In the Black-Scholes setting, the only model parameter to choose is the volatility σ, originally defined as the annualized standard deviation of logarithmic stock returns. The notion of model calibration does not exist, since after observing a trajectory of the stock price, the pricing model is completely defined. On the other hand, since the pricing model is defined by a single volatility parameter, this parameter can be reconstructed from a single option price (by inverting the Black-Scholes formula). This value is known as the implied volatility of this option.

If the real markets obeyed the Black-Scholes model, the implied volatility of all options written on the same underlying would be the same and equal to the standard deviation of returns of this underlying. However, empirical studies show that this is not the case: implied volatilities of options on the same underlying depend on their strikes and maturities (figure 5.6, left graph).

Jump-diffusion models provide an explanation of the implied volatility smile phenomenon since in these models the implied volatility is both different from the historical volatility and changes as a function of strike and maturity. Figure 5.6, right graph shows possible implied volatility patterns (as a function of strike) in the Merton jump-diffusion model.

The results of calibration of the Merton model to S&P index options are presented in figure 5.7. The calibration was carried out separately for each maturity using the routine [5] from Premia software. In this program, the vector of unknown parameters θ is found by minimizing numerically the squared norm of the

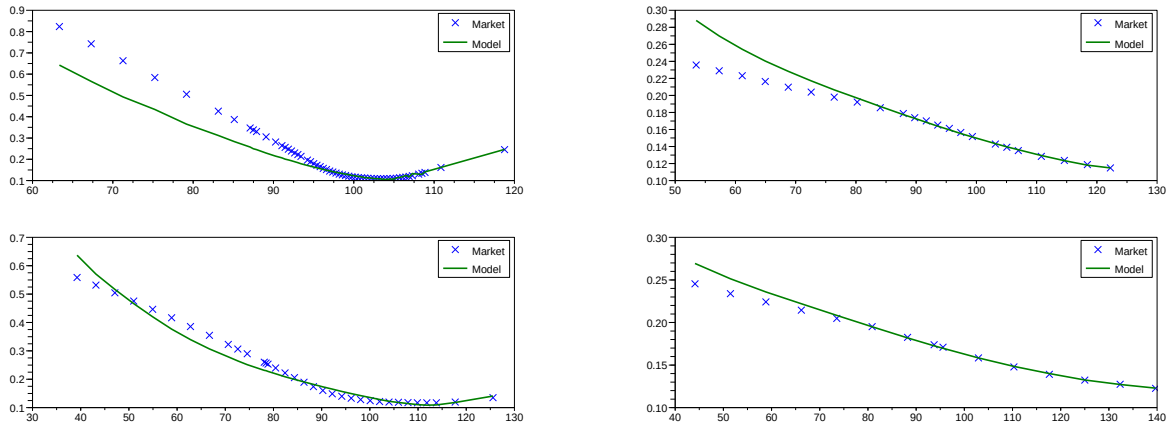


Figure 5.7: Calibration of Merton jump-diffusion model to market data separately for each maturity. Top left: maturity 1 month. Bottom left: maturity 5 months. Top right: maturity 1.5 years. Bottom right: maturity 3 years.

difference between market and model prices:

$$\theta^* = \arg\inf \|P^{obs} - P^\theta\|^2 \equiv \arg\inf \sum_{i=1}^{N} w_i (P_i^{obs} - P^\theta(T_i, K_i))^2, \tag{5.33}$$

where P^{obs} denotes the prices observed in the market and $P^\theta(T_i, K_i)$ is the Merton model price computed for parameter vector θ, maturity T_i and strike K_i. Here, the weights $w_i := \frac{1}{(P_i^{obs})^2}$ were chosen to ensure that all terms in the minimization functional are of the same order of magnitude. The model prices were computed simultaneously for all strikes present in the data using the FFT-based algorithm described in section 5.8. The functional in (5.33) was then minimized using a quasi-newton method (LBFGS-B described in [10]). In the case of Merton model, the calibration functional is sufficiently well behaved, and can be minimized using this convex optimization algorithm. In more complex jump-diffusion models, in particular, when no parametric shape of the Lévy measure is assumed, a penalty term must be added to the distance functional in (5.33) to ensure convergence and stability. This procedure is described in detail in [14, 15, 57].

The calibration for each individual maturity is quite good, however, although the options of different maturities correspond to the same trading day and the same underlying, the parameter values for each maturity are different, as seen from table 5.1. In particular, the behavior for short (1 to 5 months) and long (1 to 3 years) maturities is qualitatively different, and for longer maturities the mean jump size tends to increase while the jump intensity decreases with the length of the holding period.

Figure 5.8 shows the result of simultaneous calibration of Merton model to options of 4 different maturities, ranging from 1 month to 3 years. As we see, the calibration error is much bigger than in figure 5.7. This happens because for processes with independent and stationary increments (and the log-price in Merton model is an example of such process), the law of the entire process is completely determined by its law at any given time t (this follows from the Lévy-Khintchine formula — equation 5.10). If we have calibrated the model parameters for a single maturity T, this fixes completely the risk-neutral stock price distribution for all other maturities. A special kind of maturity dependence is therefore hard-wired into every Lévy jump diffusion model, and table 5.1 shows that it does not always correspond to the term structures of market option prices.

To calibrate a jump-diffusion model to options of several maturities at the same time, the model must have a sufficient number of degrees of freedom to reproduce different term structures. This is possible for example in the Bates model (5.34), where the smile for short maturities is explained by the presence of jumps whereas the smile for longer maturities and the term structure of implied volatility is taken into account using the stochastic volatility process. Figure 5.9 shows the calibration of the Bates model to the same data set as above. As we see, the calibration quality has improved and is now almost as good as when each maturity was calibrated separately. The calibration was once again carried out using the tool [5] from Premia.

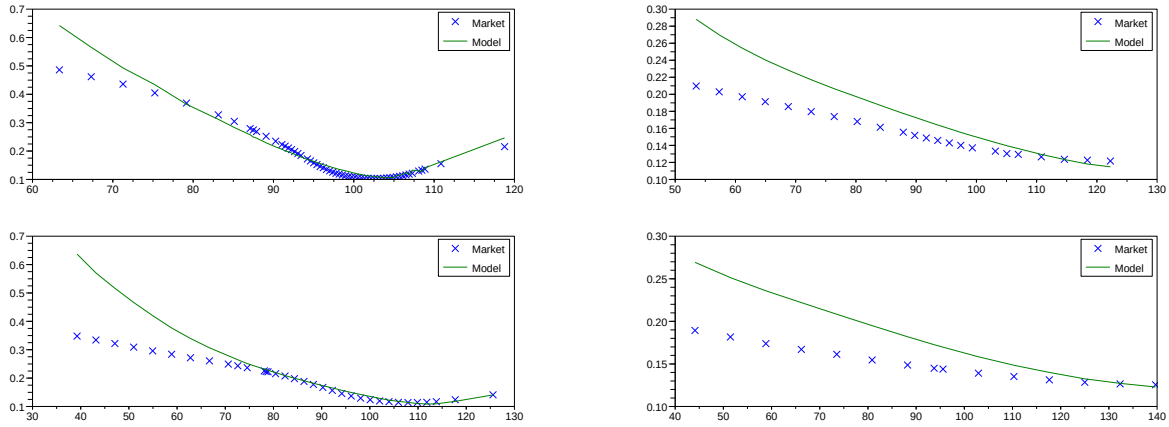


Figure 5.8: Calibration of Merton jump-diffusion model simultaneously to 4 maturities. Calibrated parameter values: $\sigma = 9.0\%$, $\lambda = 0.39$, jump mean -0.12 and jump standard deviation 0.15. Top left: maturity 1 month. Bottom left: maturity 5 months. Top right: maturity 1.5 years. Bottom right: maturity 3 years.

Maturity	$\boldsymbol{\sigma}$	$\boldsymbol{\lambda}$	**jump mean**	**jump std. dev.**
1 month	9.5%	0.097	-1.00	0.71
2 months	9.3%	0.086	-0.99	0.63
5 months	10.8%	0.050	-0.59	0.41
11 months	7.1%	0.70	-0.13	0.11
17 months	8.2%	0.29	-0.25	0.12
23 months	8.2%	0.26	-0.27	0.15
35 months	8.8%	0.16	-0.38	0.19

Table 5.1: Calibrated Merton model parameters for different times to maturity.

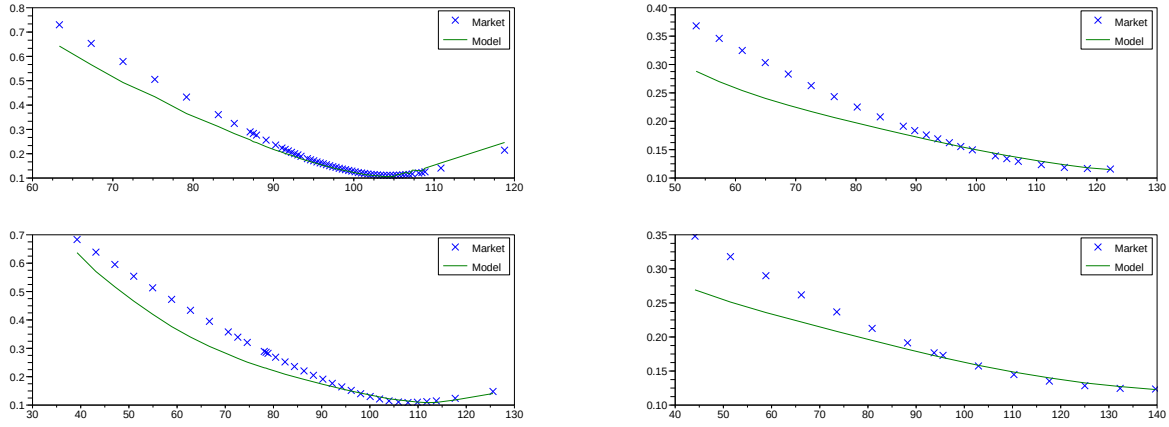


Figure 5.9: Calibration of the Bates stochastic volatility jump-diffusion model simultaneously to 4 maturities. Top left: maturity 1 month. Bottom left: maturity 5 months. Top right: maturity 1.5 years. Bottom right: maturity 3 years. Calibrated parameters (see equation (5.34)): initial volatility $\sqrt{V_0} = 12.4\%$, rate of volatility mean reversion $\xi = 3.72$, long-run volatility $\sqrt{\eta} = 11.8\%$, volatility of volatility $\theta = 0.501$, correlation $\rho = -48.8\%$, jump intensity $\lambda = 0.038$, mean jump size -1.14, jump standard deviation 0.73.

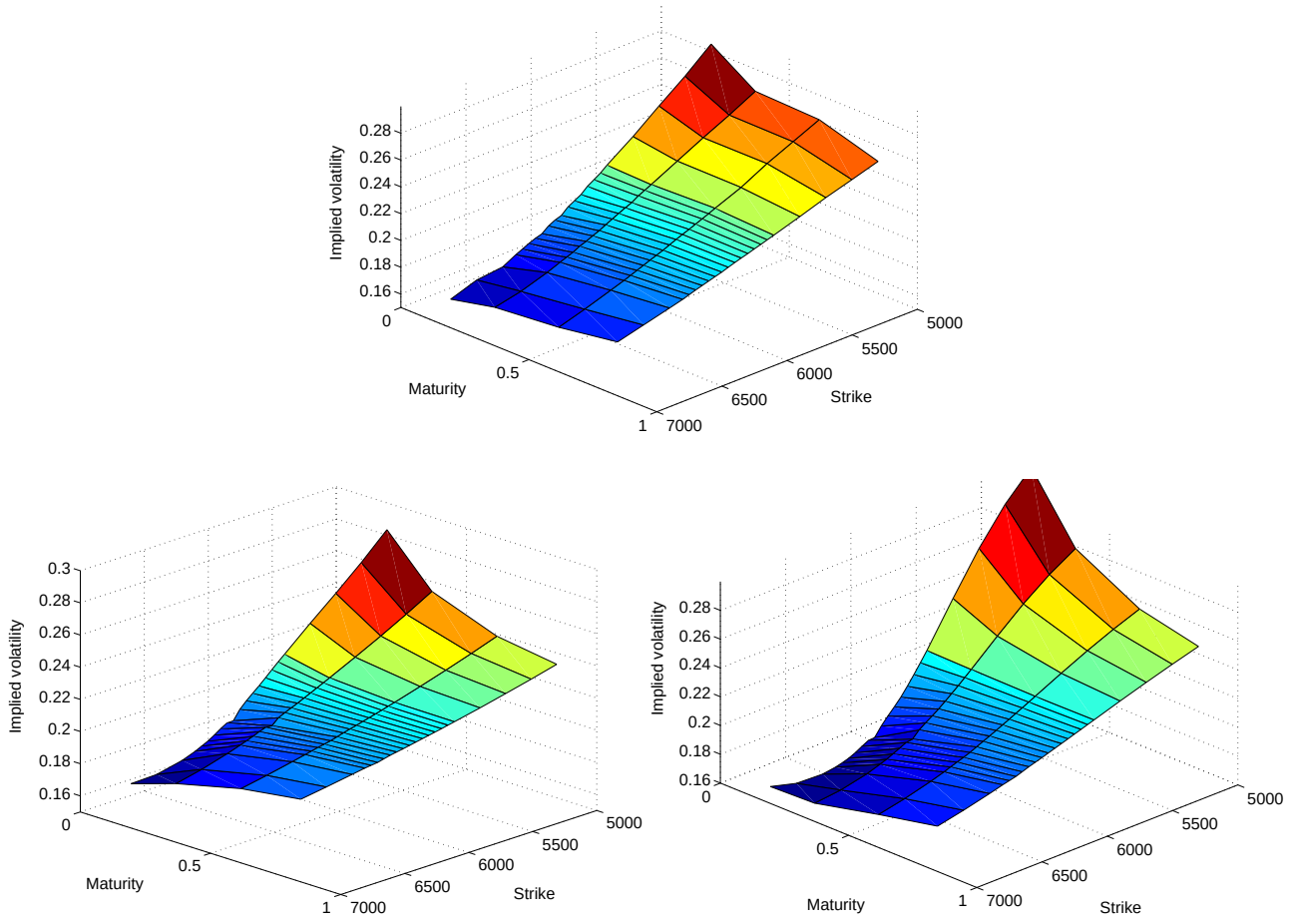


Figure 5.10: Top: Market implied volatility surface. Bottom left: implied volatility surface in an exponential Lévy model, calibrated to market prices of the first maturity. Bottom right: implied volatility surface in an exponential Lévy model, calibrated to market prices of the last maturity.

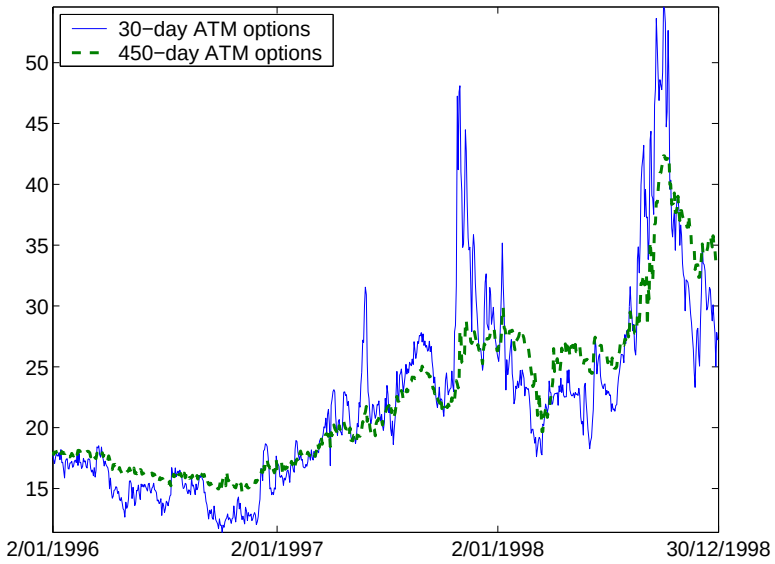


Figure 5.11: Implied volatility of at the money European options on CAC40 index.

5.10 Limites et extensions des processus de Lévy

Despite the fact that Lévy processes reproduce the implied volatility smile for a single maturity quite well, when it comes to calibrating several maturities at the same time, the calibration by Lévy processes becomes much less precise. This is clearly seen from the three graphs of Figure 5.10. The top graph shows the market implied volatilities for four maturities and different strikes. The bottom left graphs depicts implied volatilities, computed in an exponential Lévy model calibrated using a nonparametric algorithm to the first maturity present in the market data. One can see that while the calibration quality is acceptable for the first maturity, it quickly deteriorates as the time to maturity increases: the smile in an exponential Lévy model flattens too fast. The same effect can be observed in the bottom right graph: here, the model was calibrated to the last maturity, present in the data. As a result, the calibration quality is poor for the first maturity: the smile in an exponential Lévy model is more pronounced and its shape does not resemble that of the market.

It is difficult to calibrate an exponential Lévy model to options of several maturities because due to independence and stationarity of their increments, Lévy processes have a very rigid term structure of cumulants. In particular, the skewness of a Lévy process is proportional to the inverse square root of time and the excess kurtosis is inversely proportional to time [48]. A number of empirical studies have compared the term structure of skewness and kurtosis implied in market option prices to the skewness and kurtosis of Lévy processes. Bates [4], after an empirical study of implicit kurtosis in $\$/DM$ exchange rate options concludes that "while implicit excess kurtosis does tend to increase as option maturity shrinks, ..., the magnitude of maturity effects is not as large as predicted [by a Lévy model]". For stock index options, Madan and Konikov [48] report even more surprising results: both implied skewness and kurtosis actually decrease as the length of the holding period becomes smaller. It should be mentioned, however, that implied skewness/kurtosis cannot be computed from a finite number of option prices with high precision.

A second major difficulty arising while trying to calibrate an exponential Lévy model is the time evolution of the smile. Exponential Lévy models belong to the class of so called "sticky moneyness" models, meaning that in an exponential Lévy model, the implied volatility of an option with given moneyness (strike price to spot ratio) does not depend on time. This can be seen from the following simple argument. In an exponential Lévy model Q, the implied volatility σ of a call option with moneyness m, expiring in τ years, satisfies:

$$e^{-r\tau}E^Q[(S_te^{r\tau+X_\tau}-mS_t)^+|\mathcal{F}_t]=e^{-r\tau}E[(S_te^{r\tau+\sigma W_\tau-\frac{\sigma^2}{2}\tau}-mS_t)^+|\mathcal{F}_t]$$

Due to the independent increments property, S_t cancels out and we obtain an equation for the implied volatility σ which does not contain t or S_t. Therefore, in an exp-Lévy model this implied volatility does not depend on

date t or stock price S_t. This means that once the smile has been calibrated for a given date t, its shape is fixed for all future dates. Whether or not this is true in real markets can be tested in a model-free way by looking at the implied volatility of at the money options with the same maturity for different dates. Figure 5.11 depicts the behavior of implied volatility of two at the money options on the CAC40 index, expiring in 30 and 450 days. Since the maturities of available options are different for different dates, to obtain the implied volatility of an option with fixed maturity T for each date, we have taken two maturities, present in the data, closest to T from above and below: $T_1 \leq T$ and $T_2 > T$. The implied volatility $\Sigma(T)$ of the hypothetical option with maturity T was then interpolated using the following formula:

$$\Sigma^2(T) = \Sigma^2(T_1)\frac{T_2 - T}{T_1 - T} + \Sigma^2(T_2)\frac{T - T_1}{T_2 - T_1}.$$

As we have seen, in an exponential Lévy model the implied volatility of an option which is at the money and has fixed maturity must not depend on time or stock price. Figure 5.11 shows that in reality this is not so: both graphs are rapidly varying random functions.

This simple test shows that real markets do not have the "sticky moneyness" property: arrival of new information can alter the form of the smile. The exponential Lévy models are therefore "not random enough" to account for the time evolution of the smile. Moreover, models based on additive processes, that is, time-inhomogeneous processes with independent increments, although they perform well in calibrating the term structure of implied volatilities for a given date [13], are not likely to describe the time evolution of the smile correctly since in these models the future form of the smile is still a deterministic function of its present shape [13]. To describe the time evolution of the smile in a consistent manner, one may need to introduce additional stochastic factors (e.g. stochastic volatility).

Several models combining jumps and stochastic volatility appeared in the literature. In the Bates [3] model, one of the most popular examples of the class, an independent jump component is added to the Heston stochastic volatility model:

$$\begin{aligned} dX_t &= \mu dt + \sqrt{V_t} dW_t^X + \mathbf{d\mathbf{Z}_t}, & S_t &= S_0 e^{X_t}, \\ dV_t &= \xi(\eta - V_t)dt + \theta\sqrt{V_t} dW_t^V, & d\langle W^V, W^X\rangle_t &= \rho dt, \end{aligned} \tag{5.34}$$

where Z is a compound Poisson process with Gaussian jumps. Although X_t is no longer a Lévy process, its characteristic function is known in closed form [13, chapter 15] and the pricing and calibration procedures are similar to those used for Lévy processes.

Partie II

Outils

Chapitre 6

Régularisation des problèmes mal posés

6.1 Introduction

Ce chapitre commence la deuxième partie de ce polycopié, où au lieu d'étudier des modèles particuliers, on regarde le problème de calibration en général, pour comprendre, dans quelles conditions et avec quelles méthodes il peut être résolu. Du point de vue mathématique, les différents problèmes de calibration rencontrés en finance sont souvent des exemples de problèmes inverses mal posés. Dans ce chapitre nous alons étudier ces problèmes dans un cadre abstrait, pour pouvoir traiter les applications en calibration de modèles dans les chapitres suivants.

Problèmes directs et problèmes inverses La distinction entre un problème direct et un problème inverse n'est pas facile à définir. On pourrait dire qu'un problème direct consiste à déterminer l'effet à partir de sa cause:

- En physique: déterminer la température future à partir de la température initiale et de conductivité thermique
- En finance: calculer les prix des options à partir du prix initial du sous-jacent et de la loi d'évolution (e.g. volatilité)

Dans un problème inverse, au contraire, on cherche à identifier la cause d'un effet observé:

- En physique: identifier la conductivité thermique à partir de la température finale
- En finance: identifier la loi d'évolution du sous-jacent à partir des prix d'options

Les problèmes inverses sont beaucoup plus difficiles à résoudre que les problèmes directs correspondants car ils sont pour la plupart mal posés.

Problèmes bien posés et mal posés Un problème bien posé peut être intuitivement défini comme un problème ayant les trois propriétés suivantes (cette définition est due à Hadamard):

- Pour toutes données admissibles, une solution existe.
- Pour toutes données admissibles, elle est unique.
- La solution dépend continûment des données.

Pour rendre cette définition rigoureuse, on aura besoin de préciser la signification du mot 'admissible' ainsi que la topologie utilisée pour définir la continuité. Notons cependant qu'une forme de continuité est absolument nécessaire car les observations (mesures d'une quantité physique ou des observation des prix) contiennent toujours un certain niveau de bruit ou d'erreur, et en absence de continuité, la solution calculée pour ces données bruitées peut être très loin de la vraie solution.

L'exemple suivant montre les difficultés typiques qu'on peut rencontrer quand on cherche à résoudre un problème qui n'est pas bien posé.

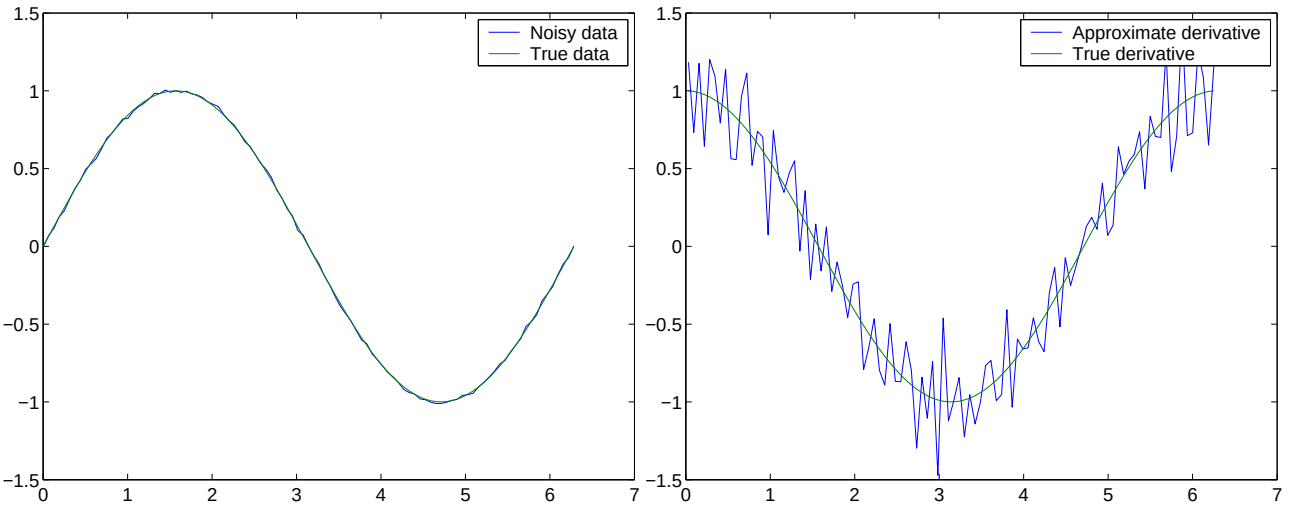


Figure 6.1: Gauche: les données exactes et bruitées. Droite: la dérivée exacte de f et la dérivée approchée de f^δ avec $h = 0.03$.

Exemple 7 (Calcul de la dérivée)**.** Supposons qu'on cherche à calculer la dérivée $f'(x)$ d'une fonction $f(x)$ à partir d'une observation bruitée $f^\delta(x)$, telle que $|f(x) - f^\delta(x)| < \delta$ pour tout x. La fonction f est supposée différentiable mais sur l'observation bruitée aucune hypothèse n'est imposée. Pour calculer la dérivée on pourrait utiliser l'approximation suivante:

$$f'(x) \approx \frac{f^\delta(x+h) - f^\delta(x-h)}{2h}$$

Pour approcher la dérivée avec une grande précision, on devrait prendre $h \to 0$. Cependant, si h est trop petit, les erreurs de données seront amplifiées:

$$\frac{f(x+h) - f(x-h)}{2h} - \frac{f^\delta(x+h) - f^\delta(x-h)}{2h} \sim O\left(\frac{\delta}{h}\right)$$

Ceci est une indication que le problème de différentiation est un problème mal posé. Pour diminuer l'erreur due aux données bruitées, on peut augmenter h, mais cela introduit une erreur de discrétisation. Si on sait que $|f'''(x)| \leq C$ alors

$$\frac{f(x+h) - f(x-h)}{2h} - f'(x) \approx \frac{h^2 C}{12}$$

L'erreur totale $\varepsilon(h) = \frac{h^2 C}{12} + \frac{\delta}{h}$ atteint son minimum pour

$$h_{opt} = \left(\frac{6\delta}{C}\right)^{1/3}, \quad \varepsilon(h_{opt}) \sim \delta^{2/3}$$

Conclusions:

- Dans un problème mal posé les erreurs d'observation sont amplifiées dans la solution de façon incontrôlable.
- Information a priori sur la solution (ici, la borne sur f''') et le niveau d'erreur (ici, la valeur de δ) peut être utilisée pour approcher la solution de façon stable.
- Une amplification d'erreurs est inévitable: la précision de l'approximation de la solution est toujours moins bonne que la précision des données d'entrée.

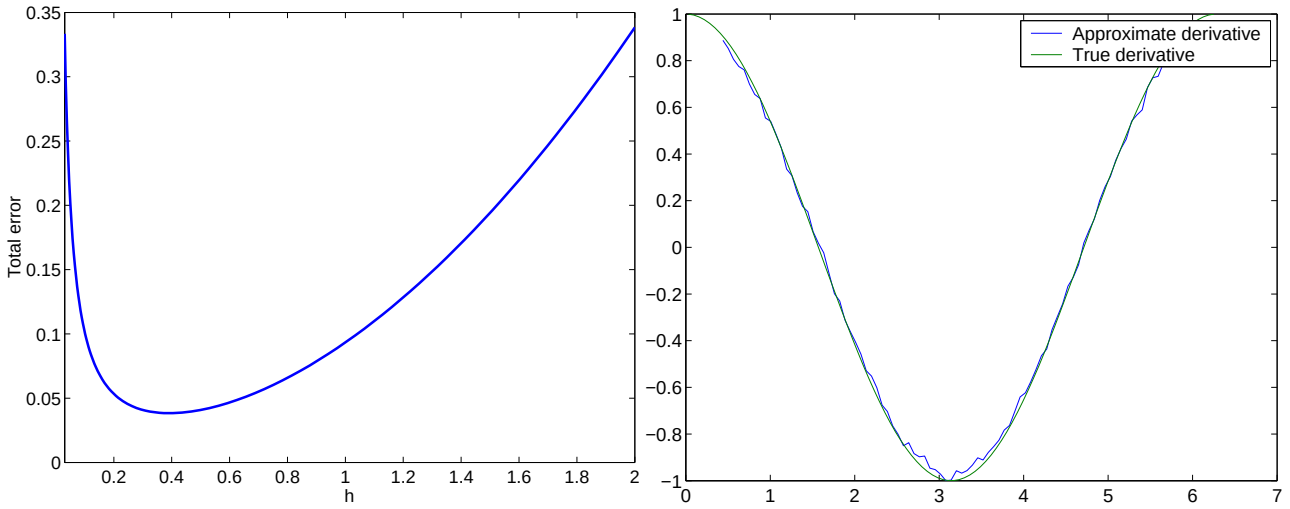


Figure 6.2: Gauche: Erreur totale (données + discrétisation) en fonction de pas de discrétisation h. Droite: la dérivée exacte de f et la dérivée approchée de f^δ avec $h_{opt} = 0.39$.

6.2 Problèmes mal posés linéaires

Dans cette section on se place dans le cadre restreint de problèmes inverses mal posés linéaires pour pouvoir définir rigoureusement un problème mal posé et étudier les techniques de régularisation de tels problèmes.

Soit $\mathcal{X}$ et $\mathcal{Y}$ deux espaces de Hilbert et $T : \mathcal{X} \to \mathcal{Y}$ un opérateur linéaire continu (rappelons que pour les opérateurs linéaires la continuité est équivalente à la bornitude). On note $\mathcal{R}(T)$ l'image de T, $\mathcal{N}(T)$ le noyau de T et T^* l'opérateur adjoint.

Le problème direct consiste a calculer Tx pour un x donné, et du fait de la continuité de T ce problème est bien posé. Le problème inverse consiste à résoudre l'équation

$$Tx = y$$

et ce problème peut être mal posé pour trois raisons suivantes:

1. Non-existence de solution: $y \notin \mathcal{R}(T)$.
2. Manque d'unicité: T n'est pas inversible, i.e., $\mathcal{N}(T) \neq \{0\}$.
3. Manque de continuité par rapport aux données: T^{-1} existe mais n'est pas continu.

En dimension fini ($\mathbb{R}^d$) tout opérateur linéaire est continu, et tout problème linéaire admettant une unique solution est bien posé, la question 3 ne se pose donc pas.

L'exemple suivant montre que ceci n'est plus vrai en dimension infini.

Exemple 8. Soit $\mathcal{X} = \mathcal{Y} = L^2([-1, 1])$ et T l'opérateur d'intégration

$$Tf(z) = \int_{-1}^{z} f(x)dx, \qquad T^{-1} = \frac{d}{dx}.$$

On montre facilement que l'opérateur T est borné et donc continu. Cependant, si on pose $f_n = \sqrt{n}(1 - n|x|)1_{|x|\leq\frac{1}{n}}$, alors $\|f_n\|^2 = \frac{2}{3}$ mais $\|T^{-1}f_n\|^2 = 2n^2$. Donc, T^{-1} n'est pas borné, ce qui implique que la différentiation est un problème mal posé.

La méthode de moindres carrées Lorsque $y \notin \mathcal{R}(T)$, en dimension fini, l'existence peut toujours être obtenue en passant à la formulation au sens de moindres carrées.

Définition 11. x est une solution moindres carrées de $Tx = y$ si

$$\|Tx - y\| = \inf\{\|Tz - y\|, z \in \mathcal{X}\}$$

Proposition 15. *Une solution au sens de moindres carrées du problème $Tx = y$ existe si et seulement si $y \in \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$. Dans ce cas elle est caractérisée par l'équation normale de Gauss:*

$$T^*Tx = T^*y. \tag{6.1}$$

Remarque 5. En dimension fini et plus généralement si $\mathcal{R}(T)$ est fermé, on a $\mathcal{R}(T) \oplus \mathcal{R}(T)^\perp = \mathcal{Y}$: tout problème linéaire a une solution au sens de moindres carrées.

Preuve. La partie "si". Dans ce le cas où $y \in \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$, on peut trouver une décomposition $y = y_1 + y_2$ avec $y_1 \in \mathcal{R}(T)$ et $y_2 \in \mathcal{R}(T)^\perp$. Alors

$$\|Tx - y\|^2 = \|Tx - y_1\|^2 + \|y_2\|^2$$

et le minimum est atteint si $Tx = y_1$ ce qui a toujours une solution. De plus, $Tx = y_1$ implique $Tx - y \in \mathcal{R}(T)^\perp$ et donc $z^*T^*(Tx - y) = 0$ pour tout $z \in \mathcal{X}$, d'où l'équation (6.1).

La partie "seulement si". Par linéarité et continuité du produit scalaire, $\mathcal{R}(T)^\perp$ est toujours un sous-espace linéaire fermé de $\mathcal{Y}$. Supposons que $y \notin \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$ et notons par y_2 la projection orthogonale de y sur $\mathcal{R}(T)^\perp$. Alors $y - y_2$ appartient à $\overline{\mathcal{R}(T)}$ mais non à $\mathcal{R}(T)$. Ceci implique que

$$\inf_x \|Tx - y + y_2\|^2 = 0,$$

mais il n'existe pas de x pour lequel $\|Tx - y + y_2\|^2 = 0$, donc, le problème n'a pas de solution au sens de moindres carrés. □

En dimension infini si $\mathcal{R}(T)$ n'est pas fermé, le problème $Tx = y$ peut ne pas avoir de solutions au sens de moindres carrées, comme le montre l'exemple suivant.

Exemple 9. Pour l'opérateur d'intégration défini dans l'exemple (8), $\mathcal{R}(T)^\perp = \{0\}$. En effet, une fonction $g \in L^2([-1, 1])$ appartient à $\mathcal{R}(T)^\perp$ si et seulement si pour toute fonction $f \in L^2([-1, 1])$,

$$\int_{-1}^{1} dx g(x) \int_{-1}^{x} f(y)dy = 0.$$

Moyennant une intégration par parties ceci est équivalent à

$$\int_{-1}^{1} dx f(x) \int_{x}^{1} g(y)dy = 0.$$

Comme ceci doit être vrai pour toute f, on en déduit $\int_x^1 g(y)dy = 0$ pour tout x et donc $g \equiv 0$. Ceci montre que $\mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$ coïncide avec $\mathcal{R}(T)$ (ensemble de fonctions différentiables) et ne coïncide pas avec $\mathcal{Y}$.

Unicité de solution Pour s'assurer de l'unicité de solution, on impose un critère supplémentaire: dans le cas linéaire on s'intéresse souvent à la *solution moindres carrées de norme minimale:*

$$x^\dagger = \arg\inf\{\|z\|, z \text{ est une solution m. c. de } Tz = y\}$$

Proposition 16. *La solution moindres carrées de norme minimale $x^\dagger$, aussi appelée* meilleure solution approchée *existe et est unique si et seulement si $y \in \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$. Elle est caractérisée par:*

$$T^*Tx^\dagger = T^*y, \qquad x^\dagger \in \mathcal{N}(T)^\perp. \tag{6.2}$$

L'application $y \mapsto x^\dagger$ est un opérateur linéaire qui s'appelle l'inverse généralisée *de T, parfois notée $T^\dagger$.*

Preuve. Par définition, la solution moindres carrées de norme minimale ne peut exister que si une solution moindres carrées existe. On peut donc supposer sans perte de généralité que $y \in \mathcal{R}(T) \oplus \mathcal{R}(T)^\perp$.

Commençons par démontrer (6.2). Supposons que $x^\dagger \neq \mathcal{N}(T)^\perp$. Comme l'opérateur T est supposé continu, $N(T)$ est un sous-espace linéaire fermé de $\mathcal{X}$ et on peut donc trouver une décomposition $x^\dagger = x_1^\dagger + x_2^\dagger$ avec $x_1^\dagger \in \mathcal{N}(T)$ et $x_2^\dagger \in \mathcal{N}(T)^\perp$. Le vecteur $x_2^\dagger$ est aussi une solution moindres carrées et sa norme est inférieure à celle de $x^\dagger$.

Pour montrer l'existence, notons que pour toute solution moindres carrées x, sa projection sur $\mathcal{N}(T)^\perp$ satisfait (6.2).

Pour l'unicité, supposons qu'il existe deux solutions x et x' de (6.2). Alors, la différence $x - x'$ satisfait. $x - x' \in N(T)^\perp$ et $T^*T(x - x') = 0$ d'où $x - x' \in N(T)$, ce qui implique $x = x'$.

La linéarité de l'application $y \to x^\dagger$ découle directement de la caractérisation (6.2). □

En dimension finie, comme tout opérateur linéaire, l'inverse généralisée $T^\dagger$ est continu et résout donc complètement un problème inverse mal posé linéaire. En dimension infinie, elle est unique mais peut ne pas être continue par rapport à y.

Exemple 10. Pour l'opérateur d'intégration défini dans l'exemple 8, comme $\mathcal{R}(T)^\perp = \{0\}$, l'inverse généralisée existe seulement pour $y \in \mathcal{R}(T)$ et coïncide donc avec l'inverse habituelle (opérateur de différentiation), qui n'est pas continue.

Exercice 26. On cherche à résoudre l'équation

$$AX = E \tag{6.3}$$

avec

$$A = \begin{pmatrix} \varepsilon & 0 & 0 \\ \varepsilon & 1 & 1 \\ \varepsilon & -1 & -1 \end{pmatrix} \qquad E = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{et} \quad X \in \mathbb{R}^3$$

1. Cette équation, a-t-elle une solution au sens strict? Calculer l'ensemble de ces solutions au sens de moindres carrés. Calculer la solution moindres carrés de norme minimale et l'inverse généralisée $A^\dagger$.

2. Quelle sera la solution moindres carrés de norme minimale si on remplace E par

$$E^\delta = \begin{pmatrix} 1+\delta \\ 1 \\ 0 \end{pmatrix}$$

Que peut-on dire sur la stabilité de calcul lorsque $\varepsilon \to 0$.

6.3 Régularisation de problèmes mal posés

Supposons qu'on cherche à approcher de manière stable la solution $x^\dagger = T^\dagger y$ d'un problème mal posé en utilisant les données bruitées y^δ telles que $\|y - y^\delta\| \leq \delta$. En dimension finie, il suffit de prendre $x_\delta^\dagger = T^\dagger y^\delta$, et par continuité on aura $x_\delta^\dagger \to x^\dagger$ lorsque $\delta \to 0$. En dimension infini, quand $T^\dagger$ n'est pas continu, $T^\dagger y^\delta$ peut être très loin de $T^\dagger y$. Dans ce cas, une solution stable peut être obtenu avec la méthode de *régularisation* qui consiste à remplacer l'opérateur discontinu $T^\dagger$ par une famille d'opérateurs (linéaires ou non) continus $\{R_\alpha\}_{\alpha>0}$ tels que

$$R_{\alpha(\delta)} y^\delta \to x^\dagger \quad \text{lorsque} \quad \delta \to 0$$

pour un règle de choix $\alpha(\delta)$ à préciser. L'approximation à la solution $x^\dagger$ en présence du bruit de données est alors donnée non pas par $T^\dagger y^\delta$ mais par $x_\delta^\alpha = R_{\alpha(\delta)} y^\delta$.

Etant donné qu'en dimension finie tout problème est bien posé, la méthode de régularisation la plus simple consiste a discrétiser le problème initial pour se ramener au cadre fini-dimensionnel. Cette approche a été employé dans l'exemple 7 où le rôle du paramètre α était joué par le pas de discrétisation h. Cette méthode

est connue sous le nom de régularisation par projection (sur un espace fini-dimensionnel). Cependant, bien que chaque problème fini-dimensionnel est bien posé, lorsque le niveau de bruit δ tend vers zéro, pour approcher la vraie solution avec de plus en plus de précision, on est souvent obligé d'augmenter la dimension de l'espace fini-dimensionnel de discrétisation, et cette augmentation de dimension peut empêcher la convergence de la solution discrétisée vers la vraie solution continu. Dans ce cas, on doit combiner la régularisation par projection avec d'autres méthodes de régularisation indépendantes de la dimension. Une telle méthode fait l'objet de la section suivante.

6.4 Régularisation de Tikhonov pour les problèmes linéaires

La régularisation de Tikhonov est une méthode très générale pour la régularisation de problèmes inverses mal posés linéaires et non-linéaires dans les espaces de Hilbert et Banach mais ici on ne considère que le cas linéaire. La méthode consiste à poser

$$R_\alpha y^\delta := \arg\inf_z \{\|Tz - y^\delta\|^2 + \alpha\|z\|^2\}. \tag{6.4}$$

Le terme de moindres carrées est ainsi pénalisé par la norme $\|z\|^2$ pour améliorer la stabilité de la solution. Le *paramètre de régularisation* α détermine le compromis entre la précision et la stabilité:

- La valeur de α dépend du niveau d'erreur δ dans les données: quand les données sont peu fiables, α doit être grand.
- Pour tout $\alpha > 0$ le problème régularisé est bien posé.
- Pour tout $\alpha > 0$ il y a une perte de précision par rapport à la solution moindres carrées.
- Le paramètre α doit être choisi tel que lorsque $\delta \to 0$, la solution régularisée converge vers $T^\dagger y$.

En écrivant la condition du premier ordre pour le problème de minimisation (6.4), on obtient directement

$$x_\alpha^\delta \equiv R_\alpha y^\delta = (T^*T + \alpha I)^{-1} T^* y^\delta. \tag{6.5}$$

Le résultat suivant montre que l'approche de Tikhonov permet effectivement de régulariser un problème mal posé dans le sens de la définition donnée plus haut. Pour simplifier l'exposé, nous supposons que le problème inverse admet une unique solution pour la donnée non-bruitée y (mais pas nécessairement pour la donnée bruitée y^δ) et nous nous concentrons sur la question de continuité.

Théorème 5. *Supposons que y est tel qu'il existe une unique solution $x^\dagger$ du problème $Tx = y$.*

(i) Pour tout $\alpha > 0$, l'opérateur R_α de la régularisation de Tikhonov est un opérateur linéaire continu.

(ii) Si $\alpha(\delta)$ est choisi tel que

$$\lim_{\delta\downarrow 0} \alpha(\delta) = 0 \quad et \quad \lim_{\delta\downarrow 0} \frac{\delta^2}{\alpha(\delta)} = 0, \tag{6.6}$$

alors

$$\lim_{\delta\downarrow 0} \|x_{\alpha(\delta)}^\delta - x^\dagger\| = 0.$$

(iii) Si de plus $x^\dagger \in \mathcal{R}(T^*T)$ alors la vitesse de convergence asymptotique maximale est atteinte pour $\alpha(\delta) = C\delta^{2/3}$, dans quel cas on a

$$\|x_{\alpha(\delta)}^\delta - x^\dagger\| = O(\delta^{2/3}).$$

Remarque 6. Pour l'opérateur d'intégration défini dans l'exemple 8, la condition $x^\dagger \in T^*T$ est équivalente à dire que les données non-bruitées y sont 3 fois différentiables, donc, c'est la même hypothèse qu'on a imposé dans l'exemple 7, et ceci n'est pas surprenant qu'on obtient la même vitesse de convergence $\delta^{2/3}$.

Preuve. (i) Pour tout $z \in \mathcal{Y}$, on a

$$\|T^*z\|^2 = \|(T^*T+\alpha I)(T^*T+\alpha I)^{-1}T^*z\|^2 = \alpha^2\|(T^*T+\alpha I)^{-1}T^*z\|^2 + \|T^*T(T^*T+\alpha I)^{-1}T^*z\|^2 + 2\alpha\|T(T^*T+\alpha I)^{-1}T^*z\|^2, \quad (6.7)$$

d'où $\|R_\alpha z\| \leq \frac{1}{\alpha}\|T^*z\|$, ce qui démontré la continuité de R_α (puisque T est borné).

(ii) Par l'inégalité du triangle,

$$\|x_\alpha^\delta - x^\dagger\| \leq \|x_\alpha^\delta - x_\alpha\| + \|x_\alpha - x^\dagger\|, \quad (6.8)$$

où $x_\alpha = (T^*T+\alpha I)^{-1}T^*y$. Par analogie aux problèmes d'estimation en statistique, le premier terme dans la partie droite peut être vu comme 'variance' et le deuxième comme 'biais'. On verra que le premier terme est décroissant en δ et croissant en α et que le deuxième terme est décroissant en α. La valeur optimale de α permettra ainsi d'atteindre un compromis entre ces deux termes.

Pour estimer le premier terme on utilisera le fait que pour toute fonction continue f,

$$f(T^*T)T^* = T^*f(TT^*). \quad (6.9)$$

Cette égalité est évidente pour les fonctions de type x^n et donc pour les polynômes, et on peut l'étendre à toute fonction continue en l'approchant par une suite de polynômes. En utilisant cette égalité pour la fonction $f(x) = \frac{1}{x+\alpha}$,

$$\begin{aligned}\|x_\alpha^\delta - x_\alpha\|^2 &= \|T^*(TT^*+\alpha I)^{-1}(y^\delta - y)\|^2 \\ &= \langle TT^*(TT^*+\alpha I)^{-1}(y^\delta - y), (TT^*+\alpha I)^{-1}(y^\delta - y)\rangle \\ &\leq \langle (y^\delta - y), (TT^*+\alpha I)^{-1}(y^\delta - y)\rangle \\ &\leq \delta^2\|(TT^*+\alpha I)^{-1}\| \leq \frac{\delta^2}{\alpha},\end{aligned}$$

parce que $\|(TT^*+\alpha I)^{-1}\| \leq \frac{1}{\alpha}$ par le même argument qui a été utilisé dans la première partie de cette démonstration. Ainsi, sous la condition (6.6), le premier terme dans (6.8) converge vers zéro.

Le deuxième terme dans (6.8) peut être réécrit comme

$$\|x_\alpha - x^\dagger\| = \|(T^*T+\alpha I)^{-1}T^*y - (T^*T+\alpha I)^{-1}(T^*T+\alpha I)x^\dagger\| = \|\alpha(T^*T+\alpha I)^{-1}x^\dagger\| \quad (6.10)$$

Pour démontrer sa convergence vers 0, on supposera pour simplifier l'exposé que l'opérateur T^*T a un spectre discret $\{\lambda_i\}_{i=1}^\infty$, mais les mêmes résultats peuvent être obtenu pour les opérateurs dont le spectre est continu.[1] Comme pour tout opérateur auto-adjoint dans un espace de Hilbert, les vecteurs propres $\{v_i\}_{i=1}^\infty$ associés aux valeurs propres $\{\lambda_i\}_{i=1}^\infty$ forment une base orthonormée de $\mathcal{X}$. En plus T^*T est défini positif, ce qui implique $\lambda_i \geq 0$ pour tout i, mais comme on a supposé que la solution $x^\dagger$ est unique, $\mathcal{N}(T^*T) = \{0\}$ et donc $\lambda_i > 0$ pour tout i. Soit $\{x_i\}_{i=1}^\infty$ la décomposition de $x^\dagger$ dans la base $\{v_i\}_{i=1}^\infty$. En utilisant (6.10), on a

$$\|x_\alpha - x^\dagger\|^2 = \sum_{i=1}^\infty \frac{\alpha^2 x_i^2}{(\lambda_i+\alpha)^2}.$$

Puisque $\sum_{i=1}^\infty x_i^2 = \|x^\dagger\|^2 < \infty$, pour tout $\varepsilon > 0$, on peut trouver $n \in \mathbb{N}$ tel que

$$\sum_{i=n}^\infty \frac{\alpha^2 x_i^2}{(\lambda_i+\alpha)^2} < \sum_{i=n}^\infty x_i^2 \leq \frac{\varepsilon}{2}.$$

D'un autre côté, pour tout i,

$$\frac{\alpha^2 x_i^2}{(\lambda_i+\alpha)^2} < \frac{\alpha^2 x_i^2}{\lambda_i^2},$$

[1] Tous les opérateurs compacts, notamment les opérateurs intégrales avec noyau dans $L^2(\mathcal{X},\mathcal{X})$ ont un spectre discret.

et on peut donc trouver α_0 tel que pour tout $\alpha < \alpha_0$,

$$\sum_{i=1}^{n-1} \frac{\alpha^2 x_i^2}{(\lambda_i + \alpha)^2} < \frac{\varepsilon}{2}.$$

Ceci montre la convergence du deuxième terme dans (6.8) vers zéro sous la condition (6.6).

(iii) Soit $x^\dagger = T^*Tz$ pour un $z \in \mathcal{X}$. Alors en utilisant (6.10), on a

$$\|x_\alpha - x^\dagger\| = \alpha \|(T^*T + \alpha I)^{-1} T^*Tz\| \leq \alpha \|z\|,$$

d'où

$$\|x_\alpha^\delta - x^\dagger\| \leq \frac{\delta}{\sqrt{\alpha}} + C\alpha.$$

pour une constante $C > 0$. On en déduit immédiatement la valeur asymptotiquement optimale de $\alpha(\delta)$ ainsi que la vitesse de convergence associée. □

Exemple 11 (Régularisation de Tikhonov pour l'opérateur de différentiation)**.** Pour régulariser le problème de différentiation d'une fonction $g \in L^2([-1,1])$ suivant l'approche de Tikhonov, on devrait résoudre le problème suivant:

$$g = \arg \inf_{f \in L^2([-1,1])} \{\|Tf - g\|_2^2 + \alpha \|f\|_2^2\},$$

où T est l'opérateur d'intégration défini dans l'exemple 8. Il est plus simple de rechercher la fonction $u = Tf$ au lieu de f et minimiser ainsi

$$J(u) = \{\|u - g\|_2^2 + \alpha \|u'\|_2^2\}.$$

sous la contrainte $u(-1) = 0$. La condition de 1$^{\text{er}}$ ordre pour ce problème est

$$\frac{dJ(u + \varepsilon h)}{d\varepsilon}|_{\varepsilon=0} = 0$$

pour toute fonction test h avec $h(-1) = 0$. Ceci implique

$$\int_{-1}^{1} h(x)(u(x) - g(x))dx - \int_{-1}^{1} h(x)u''(x)dx + \alpha h(1)u'(1) = 0.$$

La fonction u est donc solution de

$$\alpha u''(x) - u(x) = -g(x) \tag{6.11}$$

avec conditions aux bords $u(-1) = 0$ et $u'(1) = 0$. La forme générale de la solution de (6.11) est

$$u(x) = u_0(x) + Ae^{-x/\sqrt{\alpha}} + Be^{x/\sqrt{\alpha}},$$

où u_0 est une solution particulière de (6.11). On montre par différentiation qu'une forme possible de u_0 est

$$u_0(x) = \int_{-1}^{1} \frac{g(z)}{2\sqrt{\alpha}} e^{-\frac{|x-z|}{\sqrt{\alpha}}} dz.$$

En imposant les conditions aux bords, on trouve

$$u(x) = \int_{-1}^{1} \frac{g(z)}{\sqrt{\alpha}} \frac{\sinh\left(\frac{x+z}{\sqrt{\alpha}}\right) + e^{2/\sqrt{\alpha}} \sinh\left(\frac{x-z}{\sqrt{\alpha}}\right)}{e^{2/\sqrt{\alpha}} + e^{-2/\sqrt{\alpha}}} + \int_{-1}^{1} \frac{g(z)}{2\sqrt{\alpha}} e^{-\frac{|x-z|}{\sqrt{\alpha}}} dz.$$

Finalement, en revenant à $f = u'$, on a

$$f(x) = \int_{-1}^{1} \frac{g(z)}{\alpha} \frac{\cosh\left(\frac{x+z}{\sqrt{\alpha}}\right) + e^{2/\sqrt{\alpha}} \cosh\left(\frac{x-z}{\sqrt{\alpha}}\right)}{e^{2/\sqrt{\alpha}} + e^{-2/\sqrt{\alpha}}} + \int_{-1}^{1} \frac{g(z)}{\alpha} e^{-\frac{|x-z|}{\sqrt{\alpha}}} \operatorname{sign}(x - z) dz.$$

Le premier terme est dû aux conditions aux bords; il disparaîtrait si au lieu de $[-1, 1]$ on travaillait sur un intervalle infini. Le terme important est le deuxième terme: il montre que la régularisation de Tikhonov pour l'opérateur de différentiation consiste à lisser la fonction à dériver avec un noyau exponentiel et appliquer l'opérateur au résultat.

Exercice 27. On se place dans le cadre de l'exercice 26.

1. Calculer la solution de ce problème pour la donnée E^δ en utilisant la régularisation de Tikhonov avec paramètre de régularisation α donnée par

$$X_\alpha^\delta = \arg\inf_X \{\|AX - E^\delta\|^2 + \alpha\|X\|^2\}$$

pour un $\alpha > 0$ fixé, que peut-on dire sur la stabilité de calcul lorsque $\varepsilon \to 0$.

2. Supposons que $\alpha(\delta)$ est choisi tel que $\alpha(\delta) \to 0$ quand $\delta \to 0$. Montrer que $X_{\alpha(\delta)}^\delta \to X^\dagger$ lorsque $\delta \to 0$, où $X^\dagger$ est la solution de norme minimale de (B.1). Pourquoi, dans ce cas on n'a pas besoin d'imposer une condition supplémentaire (voir le cours) sur $\alpha(\delta)$ pour avoir la convergence?

Choix du paramètre de régularisation Les méthodes a priori utilisent l'information sur le niveau d'erreur δ et sur l'opérateur T ou sur la (régularité de) solution. Par exemple, si on sait que $x^\dagger \in \mathcal{R}(T^*T)$, on peut poser $\alpha = \delta^{\frac{2}{3}}$ (méthode asymptotiquement optimal) .

Les méthodes a posteriori: utilisent aussi les données y^δ. La méthode a posteriori la plus utilisée est connue sous le nom du principe de discrépance (divergence). La fonction de discrépance d'un problème mal posé pour une donnée y^δ fixée est définie par

$$\varepsilon(\alpha) = \|Tx_\alpha^\delta - y^\delta\|, \tag{6.12}$$

où x_α^δ est donné par (6.5). Le principe consiste alors à choisir α maximal pour la discrépance est encore acceptable, i.e., légèrement supérieure à l'erreur de données.

$$\alpha_{opt} = \max\{\alpha : \varepsilon(\alpha) = c\delta\},$$

où $c > 1$ est une constante proche de 1. On est obligé de prendre une constante strictement supérieure à 1 pour s'assurer que α est toujours positif. Le résultat suivant montre que cette procédure est bien définie (si $\delta \geq \|y^\delta\|$ il faut prendre $\alpha = \infty$ ce qui correspond à $x_\alpha^\delta = 0$).

Proposition 17. *La fonction de discrépance* (6.12) *est une fonction croissante continue qui satisfait*

$$\lim_{\alpha\downarrow 0} \varepsilon(\alpha) \leq \delta \qquad \lim_{\alpha\uparrow\infty} \varepsilon(\alpha) = \|y^\delta\|$$

Preuve. En utilisant l'identité (6.9), on obtient facilement

$$\varepsilon(\alpha) = \alpha\|(TT^* + \alpha I)^{-1} y^\delta\|,$$

d'où la continuité et l'égalité $\lim_{\alpha\uparrow\infty} \varepsilon(\alpha) = \|y^\delta\|$. D'un autre côté,

$$\varepsilon(\alpha) \leq \alpha\|(TT^* + \alpha I)^{-1}(y - y^\delta)\| + \|Tx_\alpha - y\|.$$

La norme du premier terme est inférieure à δ, et dans la preuve du théorème 5 nous avons vu que $\|x_\alpha - x^\dagger\| \to 0$ lorsque $\alpha \to 0$ d'où $\|Tx_\alpha - y\| \to 0$. □

La régularisation de Tikhonov avec le paramètre α choisi selon le principe de discrépance converge toujours et permet d'atteindre la vitesse optimale si $x^\dagger$ satisfait des conditions de régularité nécessaires, e.g. si $x^\dagger \in \mathcal{R}(T^*T)$ alors la vitesse est la même que dans théorème 5. Le grand avantage du principe de discrépance est que la connaissance de la régularité de $x^\dagger$ n'est pas nécessaire pour le choix de paramètre α.

6.5 Régularisation du problème de reconstruction de la courbe de taux d'intérêt

La courbe de taux zéro-coupons Le taux zéro-coupon $R(T)$ de maturité T (vu de 0) est défini par

$$B(0,T) = e^{-TR(T)},$$

où $B(0,T)$ est le prix en 0 d'un zéro-coupon (obligation qui paie 1 euro en T).

Les zéro-coupon ne sont pas cotés, mais les taux LIBOR, les taux swap et les prix d'obligations permettent de les calculer pour 20-30 échéances allant d'un jour (taux EONIA) à 30 ans. En particulier, les courbes de taux zéro-coupon sont disponibles chez les fournisseurs de données tels que Bloomberg ou Datastream. Cependant, il est souvent nécessaire de connaître les taux zéro pour les dates qui ne sont pas présentes dans les données de marché. On doit alors résoudre le problème suivant:

$$\begin{array}{l}\text{Partant des taux zéro } R_1,\dots,R_N, \text{ correspondant aux}\\ \text{échéances } T_1,\dots,T_N, \text{ trouver } R(T) \text{ pour tout } T.\end{array} \qquad (ZC)$$

Une alternative consiste à utiliser directement les prix des obligations et les taux swap, sans passer par les prix de zéro-coupons. De façon générale, le prix d'une obligation de nominal N versant des coupons $c_i, i=1,\dots,K$ aux dates $T_1,\dots,T_K$ est donné par

$$P=\sum_{i=1}^{K} c_i B(T_i) + NB(T_K).$$

Notant P_i le prix de i-ième obligation et C_j^i son payoff à la date T_j (coupon ou remboursement de nominal), le problème de reconstruction de la courbe de taux s'écrit sous la forme

$$\begin{array}{l}\text{Trouver } B(T) \text{ pour tout } T \text{ sous contrainte}\\ \displaystyle\sum_{j=1}^{K} C_j^i B(T_j) = P_i, \quad i=1,\dots,N,\\ \text{où } N \text{ est le nombre total d'obligations.}\end{array} \qquad (OBLIG)$$

L'information sur les taux swap peut aussi s'écrire sous cette forme. Dans la suite, pour simplifier l'exposé, on se limitera au problème de reconstruction de la courbe de taux sous la forme (ZC). Les principaux objectifs d'une méthode de reconstruction de la courbe de taux sont les suivantes:

- Régularité de la courbe de taux zéro-coupon et de la courbe de taux forward instantanés $f(T)$ donnée par

$$f(T)=(TR(T))'.$$

- Précision: en général on demande que les écarts entre les données de marché et la courbe interpolée n'excèdent pas un point de base ou un niveau similaire.

- Stabilité des résultats par rapport aux perturbations des données.

6.5.1 Méthodes classiques

Méthodes classiques d'interpolation. Une méthode simple pour résoudre le problème (ZC) consiste à interpoler les points de marché par une méthode classique d'interpolation, par exemple, par splines cubiques. La méthode de splines cubiques consiste à interpoler les données $R_1,\dots,R_N$ connues aux points $T_1<\dots<T_N$ par une courbe $R(T)$ telle que

- $R(T)$ est un polynôme de degré 3 sur $[T_i,T_{i+1}]$, $i=1,\dots,N-1$

- R'' est continu sur $[T_1,T_N]$;

- $R''(T_1)=R''(T_N)=0$ (un spline cubique qui satisfait cette dernière condition s'appelle spline cubique naturel).

Comme on voit sur la figure 6.3, graphique de gauche, les méthodes classiques d'interpolation ne permettent pas de construire des courbes zéro et forward de régularité suffisante.

Méthodes paramétriques. Dans les méthodes paramétriques on choisit une forme fonctionnelle paramétrique pour la courbe de taux forward ou pour la courbe de taux zéro: $R(T) = R_\theta(T)$ où θ est le vecteur de paramètres. Le paramètres sont ensuite calibrés par la méthode de moindres carrées:

$$\theta^* = \arg\inf_\theta \sum_{i=1}^N (R_i - R_\theta(T_i))^2.$$

Les exemples le plus couramment utilisés de cette approche sont la méthode de Nelson-Siegel où on écrit la courbe de taux forward sous la forme

$$f^{NS}(t) = \theta_0 + \theta_1 e^{-t/\tau_1} + \theta_2 \frac{t}{\tau_1} e^{-t/\tau_1}$$

et la méthode de Swensson où

$$f^S(t) = \theta_0 + \theta_1 e^{-t/\tau_1} + \theta_2 \frac{t}{\tau_1} e^{-t/\tau_1} + \theta_3 \frac{t}{\tau_2} e^{-t/\tau_2}.$$

Notons que ceci implique pour le modèle de Nelson-Siegel

$$R^{NS}(t) = \theta_0 + \theta_1 \frac{\tau_1}{t}(1 - e^{-t/\tau_1}) - \theta_2 e^{-t/\tau_1} + \theta_2 \frac{\tau_1}{t}(1 - e^{-t/\tau_1})$$

et pour le modèle de Swensson

$$R^S(t) = \theta_0 + \theta_1 \frac{\tau_1}{t}(1 - e^{-t/\tau_1}) - \theta_2 e^{-t/\tau_1} + \theta_2 \frac{\tau_1}{t}(1 - e^{-t/\tau_1}) \\ - \theta_3 e^{-t/\tau_2} + \theta_3 \frac{\tau_2}{t}(1 - e^{-t/\tau_2}).$$

Dans les deux modèles, la courbe de taux zéro a donc la forme

$$R(t) = \sum_{j=0}^M \theta_j \alpha_j(t, \tau),$$

où le vecteur τ contient les paramètres 'non-linéaires'. Les valeurs de $\theta_0, \ldots, \theta_M$ peuvent dont être calculés analytiquement: si on note $A_{ij}(\tau) = \alpha_j(T_i, \tau)$ et R le vecteur de données de marché, on a, pour τ fixé,

$$\arg\min_\theta \|R - A(\tau)\theta\|^2 = (A(\tau)^* A(\tau))^{-1} A(\tau)^* R.$$
$$\min_\theta \|R - A(\tau)\theta\|^2 = R^* R - R^* A(\tau)(A(\tau)^* A(\tau))^{-1} A(\tau)^* R.$$

Les valeurs optimales de paramètres τ et θ sont donc données par

$$\tau^{opt} = \arg\sup_\tau R^* A(\tau)(A(\tau)^* A(\tau))^{-1} A(\tau)^* R,$$
$$\theta^{opt} = (A^*(\tau^{opt}) A(\tau^{opt}))^{-1} A^*(\tau^{opt}) R.$$

Dans le modèle de Nelson-Siegel, on doit ainsi résoudre numériquement seulement un problème de minimisation unidimensionnel.

La figure 6.3, graphique de droite, montre que la précision de méthodes paramétriques peut parfois se révéler insuffisante.

6.5.2 Méthode de régularisation

Dans cette section nous étudions une méthode non-paramétrique de reconstruction de la courbe de taux d'intérêt, inspirée de la régularisation de Tikhonov et proposé dans un document de travail par H. Berestycki et J. Busca. L'idée principale consiste à introduire un terme de lissage de la courbe:

$$R^* = \arg\inf_R \left\{ \sum_{i=1}^N (R(T_i) - R_i)^2 + \alpha \int_0^T (tR''(t))^2 dt \right\} \tag{6.13}$$

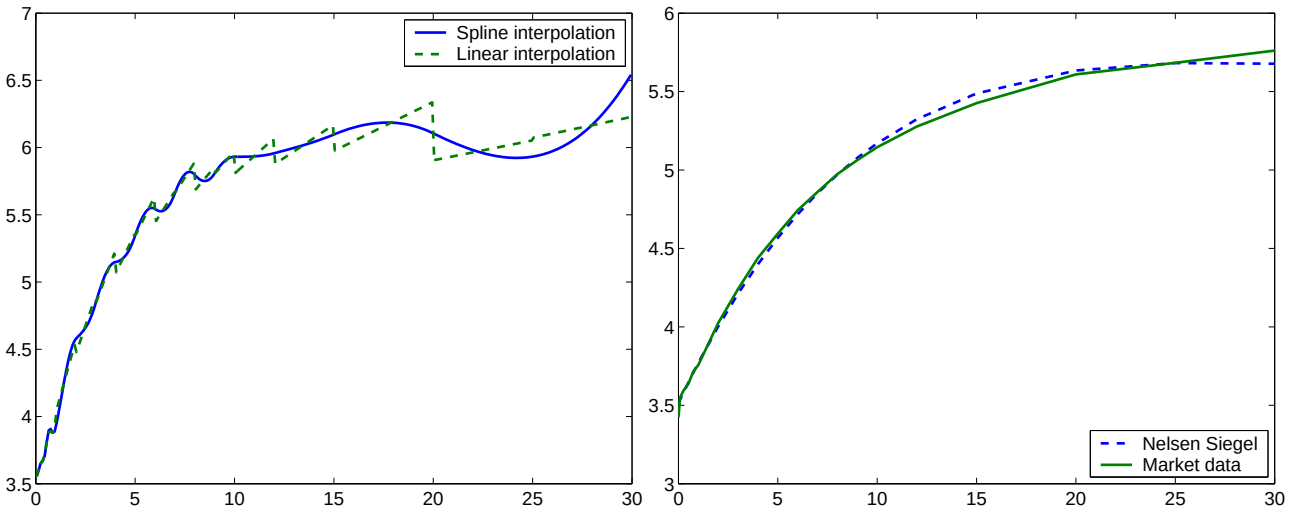


Figure 6.3: Gauche: courbe forward calculée avec taux zéro-coupon interpolés. Droite courbe zéro-coupon obtenue par la méthode de Nelson Siegel. L'écart maximal est de 9 points.

Ici, α est un paramètre de réglage qui détermine le compromis entre la précision (α petit) et la régularité de la courbe (α grand).[2]

Techniquement, le deuxième terme dans la fonctionnelle ci-dessus n'est pas une norme, mais pour $N \geq 2$ le problème (6.13) peut être réécrit pour se ramener au cadre de régularisation de Tikhonov:

$$R^* = R_1 + (R_N - R_1)\frac{T - T_1}{T_N - T_1} + \arg\inf_{\tilde{R}} \tilde{J}(\tilde{R})$$

$$\begin{aligned} \text{avec} \quad \tilde{J}(\tilde{R}) &= \sum_{i=2}^{N-1} (\tilde{R}(T_i) - \tilde{R}_i)^2 \\ &\quad + (\tilde{R}(T_1))^2 + (\tilde{R}(T_N))^2 + \alpha \int_0^T (t\tilde{R}''(t))^2 dt \\ &= \sum_{i=2}^{N-1} (\tilde{R}(T_i) - \tilde{R}_i)^2 + \alpha \|\tilde{R}\|^2 \end{aligned}$$

$$\text{où} \quad \tilde{R}_i = R_i - R_1 - (R_N - R_1)\frac{T - T_1}{T_N - T_1}$$

$$\text{et} \quad \|R\| := \sqrt{\frac{1}{\alpha}(R(T_1))^2 + \frac{1}{\alpha}(R(T_N))^2 + \int_0^T (tR''(t))^2 dt}$$

Il est facile de vérifier que $\|R\|$ a toutes les propriétés de la norme:

1. Positivité: $\|R\| \geq 0$ et $\|R\| = 0$ si et seulement si $R \equiv 0$.
2. Homogénéité: $\|kR\| = |k| \, \|R\|$ pour tout $k \in \mathbb{R}$.
3. Inégalité de triangle: $\|R_1 + R_2\| \leq \|R_1\| + \|R_2\|$ (en utilisant Cauchy-Schwartz pour la norme L^2).

On est ramené au cas déjà étudié de régularisation de Tikhonov ce qui implique l'existence de solution et la stabilité pour le problème (6.13).

[2]Cette approche est en effet équivalente à l'utilisation des *splines de lissage*: on peut démontrer que la solution optimale est donné par un polynôme d'ordre 3 par morceaux. La grande différence avec les splines cubiques considérés auparavant est que maintenant la courbe $R(T)$ ne passe pas par les points de marché, mais on peut régler la qualité de calibration en jouant sur le paramètre α, ce qui permet de trouver un équilibre entre précision et stabilité. Dans la suite on n'explorera pas cette piste car nous sommes plutôt intéressés par le lien de ce problème avec régularisation de Tikhonov qu'avec splines. De plus, l'interprétation de splines de lissage disparaît lorsque le problème est considéré sous forme (*OBLIG*), alors que l'approche de Tikhonov reste valable.

Implémentation numérique Pour résoudre ce problème numériquement, il est discrétisé sur une grille $\{x_i\}_{i=1}^M$ plus fine que la grille initiale $\{T_i\}_{i=1}^N$, c'est-à-dire, chaque point de l'ancienne grille est aussi un point de la nouvelle grille. Soit $\bar{R} = \{\bar{R}_i\}_{i=1}^M$ un vecteur de valeurs de R sur la nouvelle grille, i.e., $\bar{R}_i = R(x_i)$ Le premier terme dans (6.13) est alors approché par

$$\|S\bar{R} - R\|^2,$$

où S est la matrice $N \times M$ définie par $S_{nm} = 1$ si $x_m = T_n$ et $S_{nm} = 0$ sinon. La dérivée R'' sera calculée avec l'approximation suivante aux points $x_2, \ldots, x_{M-1}$:

$$R''(x_i) \approx R''_i = \frac{R_{i-1}(x_i - x_{i+1}) - R_i(x_{i-1} - x_{i+1}) + R_{i+1}(x_{i-1} - x_i)}{2(x_{i-1} - x_i)(x_i - x_{i+1})(x_{i-1} - x_{i+1})}$$

et l'intégrale dans (6.13) sera approché par la méthode de rectangles sur l'intervalle $[\frac{T_1+T_2}{2}, \frac{T_{N-1}+T_N}{2}]$:

$$\int_{T_1}^{T_N} (tR''(t))^2 dt \approx \sum_{i=2}^{M} (x_i R''_i)^2 \frac{x_{i-1} + x_{i+1}}{2} = \|L\bar{R}\|^2,$$

où L est une matrice $M - 2 \times M$ définie par

$$\begin{aligned}
L_{i,i} &= \frac{x_{i+1}}{2(x_{i+1} - x_i)(x_{i+2} - x_i)} \sqrt{\frac{x_i + x_{i+2}}{2}} \\
L_{i,i+1} &= -\frac{x_{i+1}}{2(x_{i+1} - x_i)(x_{i+2} - x_{i+1})} \sqrt{\frac{x_i + x_{i+2}}{2}} \\
L_{i,i+2} &= \frac{x_{i+1}}{2(x_{i+2} - x_{i+1})(x_{i+2} - x_i)} \sqrt{\frac{x_i + x_{i+2}}{2}}
\end{aligned}$$

La fonctionnelle à minimiser devient alors

$$J_{disc}(\bar{R}) = \|S\bar{R} - R\|^2 + \alpha \|L\bar{R}\|^2$$

et le résultat est solution de

$$(S^*S + \alpha L^*L)\bar{R} = S^*R. \tag{6.14}$$

Remarquons que

- S^*S est une matrice $M \times M$ diagonale avec $[S^*S]_{ii} = 1$ si x_i appartient à l'ancienne grille et 0 sinon.
- S^*R est un vecteur de dimension M, tel que $[S^*R]_i = R_j$ si $x_i = T_j$ et zéro sinon.
- La matrice L^*L est donnée par la formule

$$[L^*L]_{ij} = \sum_{k=1}^{M-2} L_{ki} L_{kj}$$

Les courbes zéro et forward calculées sont montrées sur fig. 6.4. On remarque qu'à la fois la précision de calibration et la stabilité des courbes zéro et forward sont meilleures dans ce cas que pour les deux méthodes classiques de fig. 6.3.

Pour choisir α de façon optimale, on utilise le *principe de discrépance*: on augmente α pour améliorer la stabilité tant que le niveau d'erreur reste acceptable.

α	0.01	0.1	1	10
$\max \lvert R(T_i) - R_i \rvert$	0.0035	0.0088	0.0247	0.0588

La valeur optimale pour un niveau d'erreur inférieur à 1bp est donc entre $\alpha = 0.1$ et $\alpha = 1$.

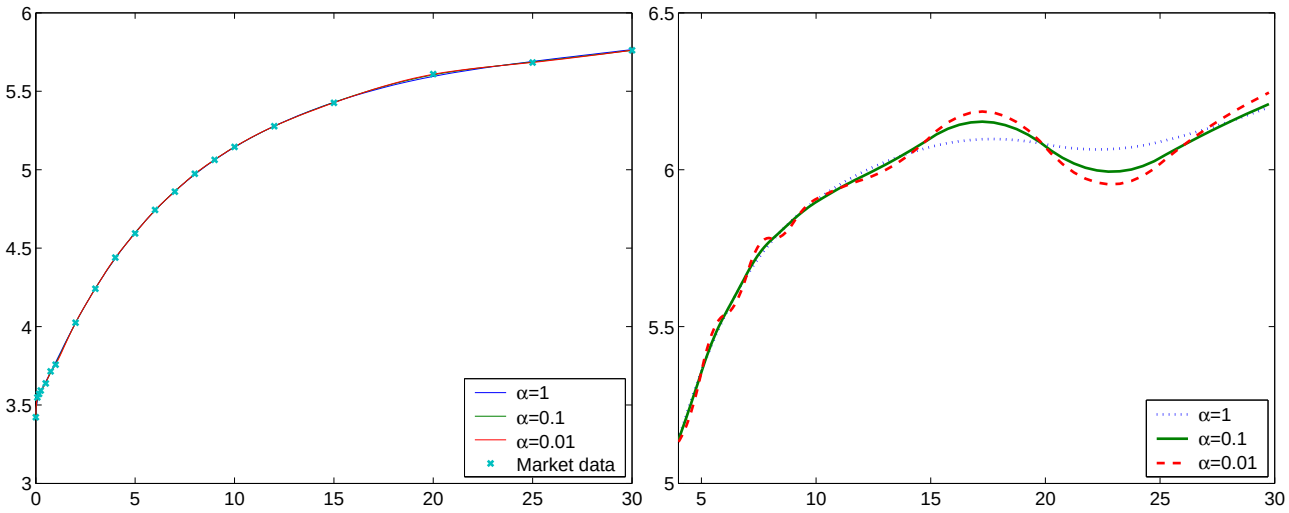


Figure 6.4: Gauche: courbes de taux zéro-coupon pour différentes valeurs du paramètre de régularisation. Droite: courbes de taux forward correspondantes.

Chapitre 7

Régularisation de problèmes de calibration I: problèmes linéaires

7.1 Reconstruction de la densité risque-neutre

Dans cette section on se place dans un marché à une période: $t_0 = 0$, $t_1 = T$ où sont cotés les actifs suivants:

- Un actif sans risque B_t: $B_0 = 1$, $B_T = e^{rT}$
- Un actif risqué S, $S_0 = 1$, S_T à la densité $p(x)$
- Des options européennes sur S de maturité T

Le théorème fondamental de valorisation des actifs implique qu'il existe une densité risque-neutre (également appelée densité des prix d'états ou SPD pour state price density): une densité $q_T \sim p_T$ telle que

$$e^{-rT} \int x q_T(x) dx = 1.$$

Le prix d'une option de payoff h satisfait

$$C_0 = e^{-rT} \int h(x) q_T(x) dx.$$

Formule de Breeden-Litzenberger Pour une option call, la fonction payoff est $h(x) = (x - K)^+$. On en déduit

$$\begin{aligned} C_{call}(K) &= e^{-rT} \int_K^\infty (x - K) q_T(x) dx \\ q_T(K) &= e^{rT} C''_{call}(K) \end{aligned}$$

La deuxième formule est connue sous le nom de formule de Breeden-Litzenberger [8]. Pour calculer q avec cette formule, il faut connaître les prix des calls pour toute valeur de strike. Or, ces prix sont connus seulement pour un ensemble fini de valeurs de K et à une fourchette bid-ask près. Pour appliquer la formule de Breeden-Litzenberger aux données de marché on doit donc interpoler les prix des options mais le problème de différentiation étant mal posé, le résultat sera très sensible aux erreurs de données et à la méthode d'interpolation. Figure 7.1 montre le résultat d'application de cette formule aux prix d'options interpolés (par splines cubiques): on observe clairement l'amplification des erreurs de données dans la solution.

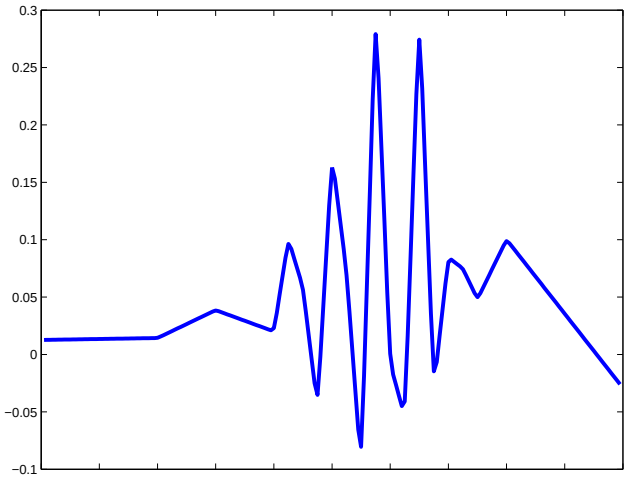


Figure 7.1: Formule de Breeden-Litzenberger avec interpolation par splines cubiques. Les erreurs dans les données sont amplifiées dans la solution.

7.1.1 Application de la régularisation de Tikhonov à la reconstruction de la SPD

Pour déterminer la densité risque-neutre q, on cherchera à régulariser l'équation

$$\int h_i(x)q_T(x)dx = C_i, \quad h_i(x) = e^{-rT}(x-K_i)^+, \quad i = 1,\dots,N$$

Avec les fonction pay-off non-bornées, la fonctionnelle $q \mapsto \int h_i(x)d(x)dx$ n'est pas bornée et la théorie du chapitre précèdent ne s'applique donc pas. Pour se ramener au cadre étudié auparavant, on remplacera les calls par les contrats butterfly (papillon) dont le payoff est à support compact

$$H_i^{but} = \frac{x-K_i}{K_{i+1}-K_i}1_{K_i<x\le K_{i+1}} + \frac{K_{i+2}-x}{K_{i+2}-K_{i+1}}1_{K_{i+1}<x\le K_{i+2}}, \qquad h_i^{but} = e^{-rT}H_i^{but}.$$

Le prix de non-arbitrage d'un contrat butterfly peut se calculer à partir des prix de marché des options call:

$$C_i^{but} = \frac{C_{i+2}-C_{i+1}}{K_{i+2}-K_{i+1}} - \frac{C_{i+1}-C_i}{K_{i+1}-K_i}.$$

Le problème initial se réécrit alors sous la forme[1]

$$\int h_i^{but}(x)q_T(x)dx = C_i^{but}, \quad i = 1,\dots,N$$

et on peut maintenant appliquer la régularisation de Tikhonov. Suivant la théorie générale, on cherche à minimiser la fonctionnelle

$$J(q) = \frac{1}{\alpha}\sum_{i=1}^{N}\left(\int h_i^{but}(x)q(x)dx - C_i^{but}\right)^2 + \int q^2(x)dx.$$

En utilisant l'identité élémentaire $\frac{b^2}{\alpha} = \sup_\lambda(2b\lambda - \alpha\lambda^2)$, le problème se transforme en

$$\inf_q \sup_{\lambda\in\mathbb{R}^N}\left\{-\alpha\|\lambda\|^2 + 2\sum_{i=1}^{N}\lambda_i\left(\int h_i^{but}(x)q(x)dx - C_i^{but}\right) + \int q^2(x)dx\right\}. \tag{7.1}$$

[1]Pour construire N prix de contrats butterfly on a en effet besoin de $N+2$ prix de calls; ce problème peut être résolu en ajoutant deux données fictives avec prix nuls et strikes suffisamment élevés.

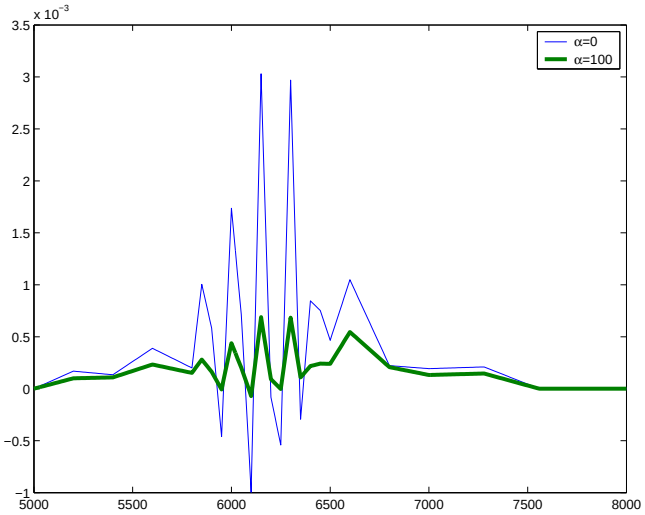


Figure 7.2: Reconstruction de q_T avec et sans régularisation

Soit

$$V(\lambda) = \inf_q \left\{ -\alpha\|\lambda\|^2 + 2\sum_{i=1}^N \lambda_i \left(\int h_i^{but}(x) q(x) dx - C_i^{but} \right) + \int q^2(x) dx \right\}. \tag{7.2}$$

On peut démontrer que si $V(\lambda)$ admet un minimum unique λ^* où elle est différentiable, alors il n'y a pas de saut de dualité et on peut interchanger le inf et le sup dans (7.1). Le inf dans (7.2) est atteint par $q(x) = -\sum_{i=1}^N \lambda_i h_i^{but}(x)$ d'où

$$\begin{aligned} V(\lambda) &= -\alpha\|\lambda\|^2 - 2\sum_{i=1}^N \lambda_i C_i^{but} - \sum_{i,j=1}^N \lambda_i \lambda_j \int h_i^{but}(x) h_j^{but}(x) dx \\ &= -\alpha\|\lambda\|^2 - 2\lambda^\perp C^{but} - \lambda^\perp H \lambda. \end{aligned}$$

où H est une matrice tridiagonale très simple à calculer:

$$H_{i,i} = e^{-2rT} \frac{K_{i+2} - K_i}{3}, \qquad H_{i,i+1} = H_{i+1,i} = e^{-2rT} \frac{K_{i+2} - K_{i+1}}{6}.$$

On voit que $V(\lambda)$ est une forme quadratique avec une matrice strictement positive. Elle admet donc un minimum unique où elle est différentiable ce qui implique que le minimum de $J(q)$ est atteint par

$$q(x) = -\sum_{i=1}^N \lambda_i^* h_i^{but}(x) \quad \text{avec} \quad \lambda^* = -(H + \alpha I)^{-1} C^{but}$$

Le résultat de l'application de cette formule aux données d'options est montré sur le graphique 7.2. Pour choisir le paramètre de régularisation α on utilise le principe de discrépance qui consiste à prendre α maximal pour lequel l'erreur peut encore être tolérée. Si ce seuil d'erreur est fixé à 0.01 le tableau suivant montre que α optimal se trouve entre 1 et 10.

α	1	10	100
Err.	0.005	0.039	0.139

Régularisation avec projection Un lecteur attentif remarquera que la fonction obtenu représentée sur le graphique 7.2 prend des valeurs négatives et n'est pas forcement d'intégrale égale à 1 car nous avons régularisé la formule de Breeden-Litzenberger sans imposer que q_T soit une densité.

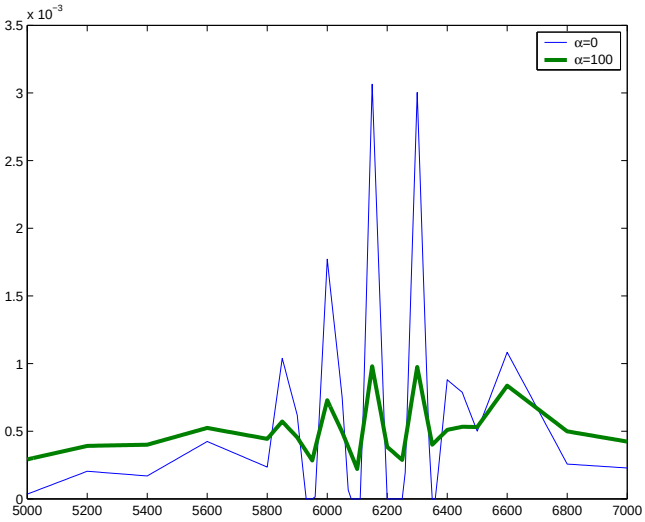


Figure 7.3: Régularisation de Tikhonov avec projection: résultats numériques.

Pour reconstruire la densité risque-neutre, il faut imposer les contraintes $q \geq 0$, $\int q(x)dx = 1$ et $e^{-rT}\int xq(x)dx = S_0$. Soit C l'ensemble convexe fermé décrit par ces contraintes. Une méthode consiste à projeter la régularisée de Tikhonov q_α sur cet ensemble à la fin du calcul.

$$q_\alpha^C = P_C q_\alpha = \arg\inf_{p\in C} \int (p(x) - q_\alpha(x))^2 dx,$$

d'où on trouve, par la méthode de multiplicateurs de Lagrange:

$$q_\alpha^C = \max(0, q_\alpha(x) - \lambda - \mu e^{-rT}x)$$

avec

$$(\lambda, \mu) = \arg\inf \int ((q_\alpha(x) - \lambda - \mu e^{-rT}x) \vee 0)^2 dx + \lambda + \mu S_0.$$

Ce deuxième problème de minimisation (en deux dimensions) doit être résolu numériquement. Le résultat de calcul (où on a imposé seulement la contrainte de positivité et la contrainte d'intégrale égale à 1 mais non la contrainte de martingale) est montre sur le graphique 7.3.

Régularisation entropique Une autre méthode pour imposer la positivité de la densité consiste à choisir une régularisation adaptée. Pour régulariser une densité on utilise souvent l'entropie relative (distance de Kullback-Leibler) de q par rapport à une mesure à priori p définie par

$$I(q|p) = \int \log\left(\frac{dq}{dp}\right)\frac{dq}{dp}dp$$

La mesure p reflète notre connaissance à priori sur la solution. La fonctionnelle à minimiser s'écrit alors

$$J(q) = \alpha^{-1}\sum_{i=1}^{N}\left(\int h_i^{but}q(x)dx - C_i^{but}\right)^2 + \int \log\left(\frac{dq}{dp}\right)\frac{dq}{dp}dp$$

Nous ne sommes plus dans le cadre de régularisation de Tikhonov (l'entropie n'est pas une norme) mais on peut toujours démontrer la convergence de cette méthode sous certaines conditions — voir [29]. La minimisation de $J(q)$ sous la contrainte de masse $q(\mathbb{R}) = 1$ donne

$$q(dx) = \frac{\exp\left(-2\sum \lambda_i h_i^{but}\right)}{\int dp \exp\left(-2\sum \lambda_i h_i^{but}\right)}p(dx)$$

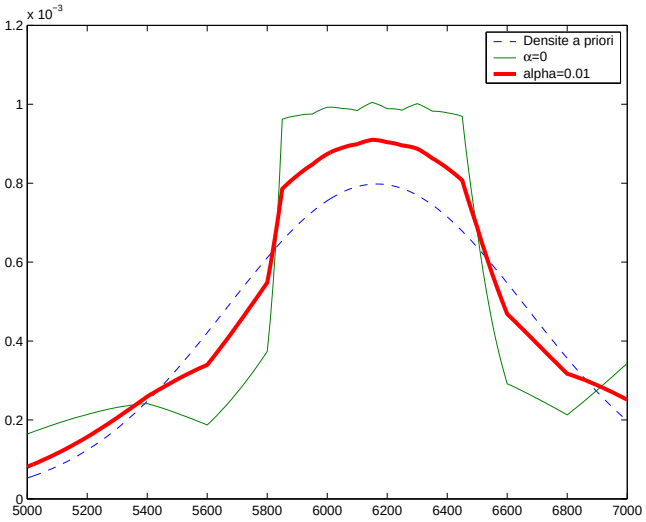


Figure 7.4: Régularisation entropique: résultats numériques.

avec

$$\lambda = \arg\inf 2\sum_{i=1}^{N} \lambda_i C_i^{but} + \alpha\|\lambda\|^2 + \log\left(\int dp e^{-2\sum \lambda_i h_i^{but}}\right).$$

On doit donc résoudre un problème de minimisation convexe N-dimensionnel (un multiplicateur de Lagrange par prix d'option). Le résultat de calcul (avec une mesure a priori gaussienne) est représenté sur le graphique 7.4. Pour plus de détails sur la régularisation entropique pour la reconstruction de la densité risque-neutre voir [1, 56].

Chapitre 8

Régularisation de problèmes de calibration II: problèmes non-linéaires

8.1 Régularisation de Tikhonov pour les problèmes non-linéaires

Dans le cas de problèmes mal posés linéaires de type $Tx = y$, les étapes de résolution sont les suivantes:

1. Pour qu'une solution existe (au moins en dimension finie), on utilise la méthode de moindres carrées:
$$\|Tx - y\| \ \rightarrow \ \min!$$
2. Pour que la solution soit unique, on passe à la solution de norme minimale (SNM) qui minimise $\|x\|$ sur l'ensemble des solutions au sens de moindres carrées.
3. Pour que la solution soit continue par rapport aux données, on applique la méthode de régularisation de Tikhonov:
$$\|Tx - y\|^2 + \alpha\|x\|^2 \ \rightarrow \ \min!$$

La méthode de Tikhonov admet une solution explicite

$$x_\alpha = (T^*T + \alpha I)^{-1}T^*y,$$

qui converge vers la SNM lorsque le niveau d'erreur $\rightarrow 0$, si le paramètre de régularisation $\alpha(\delta)$ est choisi correctement.

Dans le cas d'un problème non-linéaire $f(x) = y$ certaines de ces conclusions ne seront plus vraies:

1. La solution au sens de moindres carrées définie par
$$\|f(x) - y\| \ \rightarrow \ \min!$$
peut ne pas exister même en dimension finie (voir Fig. 8.1).
2. La solution de norme minimale doit être remplacée par x_0-solution de norme minimale qui minimise $\|x_0 - x\|$ sur l'ensemble des solutions au sens de moindres carrées, où x_0 est une valeur de référence. Cette solution n'est pas, en général, unique.
3. La régularisation de Tikhonov:
$$\|f(x) - y\|^2 + \alpha\|x - x_0\|^2 \ \rightarrow \ \min!$$
assure l'existence, la continuité et, pour α grand, l'unicité.

Cette fois, la méthode de Tikhonov n'admet pas de solution explicite, mais les solutions régularisées convergent toujours vers l'une des SNM lorsque le niveau d'erreur diminue vers zéro, si $\alpha(\delta)$ est choisi de façon appropriée.

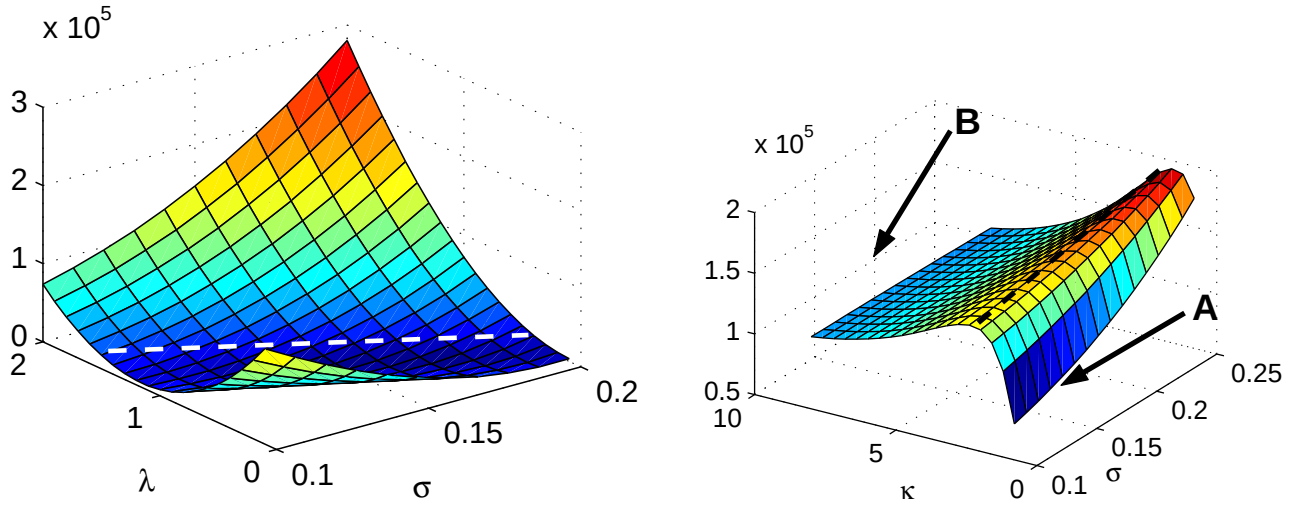


Figure 8.1: Même en dimension fini, $\|f(x) - y\|^2$ peut être non-convexe. Les graphiques illustrent la forme de la fonctionnelle à minimiser dans différents problèmes de calibration.

Rôle de régularisation: existence Toutes les solutions de

$$\inf\{\|f(x) - y\|^2 + \alpha\|x - x_0\|^2\}$$

appartiennent à l'ensemble

$$\|x - x_0\| \leq \frac{1}{\alpha}\|f(x_0) - y\|$$

En dimension finie cet ensemble est compact, et pour avoir une solution il suffit que f soit continu. En dimension infinie les hypothèses plus fines sont nécessaires: il faut que f soit *faiblement* continu.

Rôle de régularisation: continuité La définition de continuité change en présence de solutions multiples: soient deux séquences $\{x_k\}$ et $\{y_k\}$ avec $y_k \to y^\delta$ et x_k la solution du problème régularisé avec donnée y_k. Alors

- Il existe une sous-suite convergente de $\{x_k\}$.
- La limite de toute sous-suite convergente de $\{x_k\}$ est une solution du problème régularisé avec donnée y^δ.

Interpretation: si la donnée y est proche de y^*, alors chaque solution avec donnée y sera proche d'une des solutions avec donnée y^*.

Rôle de régularisation: unicité pour α grand Dans le cas où f est une fonction $\mathbb{R} \to \mathbb{R}$,

$$\frac{\partial^2}{\partial x^2}\{\|f(x) - y\|^2 + \alpha\|x - x_0\|^2\} = 2\{f'^2(x) + (f(x) - y)f''(x) + \alpha\}$$

Si f et ses dérivées sont bornées, $\|f(x) - y\|^2 + \alpha\|x - x_0\|^2$ est convexe pour α suffisamment grand.

En dimension grande ou infinie on rencontre aussi ce phénomène; le problème régularisé est plus facile à résoudre par une méthode de gradient que le problème initial.

8.2 Régularisation du problème de calcul de la volatilité locale

Dans les modèles de volatilité locale

$$\frac{dS_t}{S_t} = rdt + \sigma(t, S_t)dW_t$$

les prix d'options satisfont l'équation backward de pricing:

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2(t,S)S^2\frac{\partial^2 C}{\partial S^2} = rC, \quad C(T,S,T,K) = (S-K)^+$$

Cette équation définit l'opérateur de pricing $\sigma(*,*) \mapsto C(*,*)$. Cependant, l'opération inverse qui consiste à reconstruire $\sigma(t,S)$ pour tout t,S à partir d'un nombre fini de prix $C_M(T_i,K_i), i = 1\dots N$, est un problème mal posé.

Régularisation de Tikhonov Lagnado et Osher [44], Crepey [20] et d'autres auteurs ont utilisé la méthode de Tikhonov pour régulariser le problème de calibration de la surface de volatilité locale. Plus précisement, ces auteurs proposent de minimiser la fonctionnelle

$$J(\sigma) \equiv \sum_{i=1}^{N} w_i (C(T_i,K_i,\sigma) - C_M(T_i,K_i))^2 + \alpha \|\nabla\sigma\|_2^2, \tag{8.1}$$

$$\|\nabla\sigma\|_2^2 \equiv \int_{K_{\min}}^{K_{\max}} dK \int_{T_{\min}}^{T_{\max}} dT \left\{ \left(\frac{\partial\sigma}{\partial K}\right)^2 + \left(\frac{\partial\sigma}{\partial T}\right)^2 \right\}. \tag{8.2}$$

Techniquement, $\|\nabla\sigma\|_2$ n'est pas une norme. Pour se ramener au cadre de régularisation de Tikhonov, on pourrait prendre

$$J(\sigma) \equiv \sum_{i=1}^{N} w_i (C(T_i,K_i,\sigma) - C_M(T_i,K_i))^2 + \alpha \|\sigma\|_{H^1}^2$$

$$\text{où} \qquad \|\sigma\|_{H^1}^2 \equiv \|\nabla\sigma\|_2^2 + \int_{K_{\min}}^{K_{\max}} \int_{T_{\min}}^{T_{\max}} \sigma^2(T,K) dT\, dK$$

mais, empiriquement, cela n'améliore pas la régularisation.

Pour minimiser (8.1), on peut utiliser, par exemple, un algorithme de descente de gradient (ou la méthode de gradient conjugué — voir chapitre 9):

- Discrétiser σ sur une grille espace-temps;
- Choisir une surface initiale $\sigma^{(0)}$, par exemple, constante;
- Evaluer le gradient de $J(\sigma)$ au point $\sigma^{(0)}$ et calculer la direction de descente $\lambda^{(0)}$;
- Résoudre le problème de minimisation unidimensionnelle

$$h^* = \arg\inf_h J(\sigma^{(0)} + h\lambda^{(0)});$$

- Poser $\sigma^{(1)} := \sigma^{(0)} + h^*\lambda^{(0)}$ et passer à l'itération suivante.

Pour mettre en place cet algorithme, une étape essentielle est de calculer le gradient de J. Le gradient $\operatorname{grad} J(\sigma)(t,S)$ est défini par:

$$\langle \operatorname{grad} J(\sigma), \delta\sigma \rangle = \delta J(\sigma) \equiv \lim_{\varepsilon\to 0} \frac{J(\sigma+\varepsilon\delta\sigma) - J(\sigma)}{\varepsilon}$$

pour toute fonction test $\delta\sigma$.

Pour le deuxième terme dans (8.1), l'intégration par parties donne

$$\delta\|\nabla\sigma\|_2^2 = -2\int_{K_{min}}^{K_{\max}} dK \int_{T_{min}}^{T_{\max}} dT \left\{ \frac{\partial^2\sigma}{\partial K^2} + \frac{\partial^2\sigma}{\partial T^2} \right\} \delta\sigma,$$

d'où

$$\operatorname{grad} \|\nabla\sigma\|_2^2 = -2\left(\frac{\partial^2\sigma}{\partial K^2} + \frac{\partial^2\sigma}{\partial T^2} \right).$$

Le premier terme (noté par J_1) est plus difficile. On a

$$\delta J_1(\sigma) = 2\sum_{i=1}^{N} w_i(C(T_i,K_i,\sigma) - C_M(T_i,K_i))\delta C(T_i,K_i,\sigma)$$
$$= \left\langle 2\sum_{i=1}^{N} w_i(C(T_i,K_i,\sigma) - C_M(T_i,K_i))\delta_{(T_i,K_i)}(T,K), \delta C(T,K,\sigma)\right\rangle,$$

où le produit scalaire est ici défini par

$$\langle f,g\rangle = \int_0^{\bar T}\int_0^{\infty} fg dK dT$$

avec $\bar T$ l'horizon du temps maximal. La fonction $C(T,K,\sigma)$ satisfait l'EDP de Dupire

$$\frac{\partial C}{\partial T} = \frac{1}{2}\sigma^2(T,K)K^2\frac{\partial^2 C}{\partial K^2} - rK\frac{\partial C}{\partial K}, \quad C(t,K) = (S_t - K)^+,$$

d'où

$$\frac{\partial \delta C}{\partial T} = \frac{1}{2}\sigma^2(T,K)K^2\frac{\partial^2 \delta C}{\partial K^2} - rK\frac{\partial \delta C}{\partial K} + \sigma\delta\sigma K^2\frac{\partial^2 C}{\partial K^2},$$

ce qu'on peut écrire

$$\mathcal{L}\delta C = \sigma\delta\sigma K^2\frac{\partial^2 C}{\partial K^2}, \quad \mathcal{L} = \frac{\partial}{\partial T} - \frac{1}{2}\sigma^2(T,K)K^2\frac{\partial^2}{\partial K^2} + rK\frac{\partial}{\partial K}.$$

Soit G une fonction qui satisfait

$$\mathcal{L}^* G = 2\sum_{i=1}^{N} w_i(C(T_i,K_i,\sigma) - C_M(T_i,K_i))\delta_{(T_i,K_i)}(T,K), \quad G(\bar T,K) = 0, \tag{8.3}$$

où $\mathcal{L}^*$ est l'opérateur adjoint

$$\mathcal{L}^* f = -\frac{\partial f}{\partial T} - \frac{1}{2}\frac{\partial^2}{\partial K^2}(\sigma^2(T,K)K^2 f) - \frac{\partial}{\partial K}(rKf).$$

Alors

$$\langle \mathcal{L}^* G, \delta C(*,*,\sigma)\rangle = \langle G, \mathcal{L}\delta C(*,*,\sigma)\rangle = \langle G\sigma K^2\frac{\partial^2 C}{\partial K^2}, \delta\sigma\rangle,$$

ce qui implique

$$\operatorname{grad} J_1(\sigma) = \sigma K^2\frac{\partial^2 C}{\partial K^2}G.$$

Résolution du problème adjoint Soit

$$F(T,K) = \int_0^K dK' \int_0^{K'} dx G(T,x).$$

Alors l'équation (8.3) devient

$$-\frac{\partial F}{\partial T} - \frac{1}{2}\sigma^2K^2\frac{\partial^2 F}{\partial K^2} - rK\frac{\partial F}{\partial K} + rF$$
$$= 2\sum_{i=1}^{N} w_i(C(T_i,K_i,\sigma) - C_M(T_i,K_i))(K-K_i)^+\delta_{T_i}(T) \tag{8.4}$$

et

$$\text{grad}\, J_1(\sigma) = \sigma K^2 \frac{\partial^2 C}{\partial K^2}\frac{\partial^2 F}{\partial K^2}.$$

L'équation (8.4) est l'équation Black-Scholes standard entre les dates de maturité des options, et aux maturités, on a

$$F(T_i+,K) - F(T_i-,K) = 2\sum_{j:T_j=T_i} w_i(C(T_j,K_j,\sigma) - C_M(T_j,K_j))(K-K_j)^+.$$

Cette équation se résout par morceaux entre les dates de maturité.

Implémentation numérique Principe: calculer le gradient exact du problème discrétisé plutôt que le gradient approché du problème continu. Pour l'implémentation numérique, la volatilité locale σ doit être discrétisée suivant l'une des méthodes proposées dans la littérature:

- Différences finies: Lagnado et Osher [44].
- Arbre trinomial: Crepey [20].

Le problème est ensuite résolu par la méthode de descente de gradient dans l'espace de surfaces de volatilité discrétisées.

Points faibles de la méthode

- La fonctionnelle à minimiser $J(\sigma)$ n'est pas, en général, convexe.
- La dimension du problème de minimisation est très grande (le nombre de noeuds dans l'arbre: $10^3 \div 10^4$).
- La régularisation est la même ($\|\nabla\sigma\|$) dans les régions où la volatilité à beaucoup d'influence sur les prix et dans les régions où cette influence est très petite (là où le processus de prix a peu de chances d'arriver).

Représentation de volatilité par splines Pour réduire la dimension du problème d'optimisation, plusieurs auteurs ont propose de répresenter la volatilité par splines: [41, 12].

Splines cubiques: méthode d'interpolation d'une fonction $f(x)$ dont les valeurs f_i sont connues en un nombre fini de points $x_0,\ldots,x_n$.

$$\begin{aligned} &f(x) = a_i x^3 + b_i x^2 + c_i x + d_i \quad \text{pour} \quad x \in [x_{i-1}, x_i] \\ &f \in C^2([x_0, x_n]) \end{aligned}$$

Les coefficients du spline sont calculés à partir des équations suivantes:

$$\begin{aligned} &a_i x_{i-1}^3 + b_i x_{i-1}^2 + c_i x_{i-1} + d_i = f_{i-1}, && i = 1\ldots n \\ &a_i x_i^3 + b_i x_i^2 + c_i x_i + d_i = f_i, && i = 1\ldots n \\ &3a_i x_i^2 + 2b_i x_i + c_i = 3a_{i+1} x_i^2 + 2b_{i+1} x_i + c_{i+1}, && i = 1\ldots n-1 \\ &\text{(continuité de } f') \\ &6a_i x_i + 2b_i = 6a_{i+1} x_i + 2b_{i+1}, && i = 1\ldots n-1 \\ &\text{(continuité de } f'') \end{aligned}$$

Il restent alors deux degrés de liberté pour completement fixer les coefficients. Pour le *spline cubique naturel*, on prend $f''(x_0) = f''(x_n) = 0$. Ce spline est la fonction qui minimise $\int_{x_0}^{x_n}(f'')^2 dx$ parmi toutes les fonctions qui passent par les points $(x_i, f_i)_{i=0}^n$.

Dans l'approche de Jackson, Suli et Howison [41], $\sigma(t,S)$ est supposée 2 fois continûment différentiable en S et continue en t:

- Pour chaque t fixé, $\sigma(t,S)$ est une spline cubique avec noeuds $S_0 \ldots S_P$.
- Pour chaque $p = 0,\ldots,P$, $\sigma(t,S_p)$ est linéaire par morceaux avec noeuds $t_0,\ldots,t_Q$

- La volatilité est complétement déterminé par la matrice de coefficients $(\Sigma_{pq})_{p=0\dots P}^{q=0\dots Q}$, où $\Sigma_{pq} = \sigma(t_q, S_p)$.

Pour résoudre le problème de calibration, on minimise la fonctionnelle

$$J(\Sigma) = \sum_{i=1}^{N} w_i (C(T_i, K_i, \Sigma) - C_M(T_i, K_i))^2 + c_1 \sum_{p,q} \frac{\sigma_S^2(S_p, t_q)}{(P+1)(Q+1)} + c_2 \sum_{p,q;q\neq Q} \frac{\sigma_t^2(S_p, t_q+)}{(P+1)Q}$$

Une discrétisation par différences finies ou éléments finis est toujours nécessaire pour calculer les prix d'options mais la dimension du problème de minimisation est réduite.

Chapitre 9

Méthodes numériques pour la calibration de modèles

9.1 Algorithmes d'optimisation

9.1.1 Optimisation: rappels théoriques

Soit H un espace de Hilbert, $J : H \to \mathbb{R}$ et $K \subseteq H$. On dit que $u^* \in K$ est un minimum *global* de J sur K si

$$\forall u \in K, \quad J(u^*) \leq J(u).$$

$u^* \in K$ est un minimum *local* de J sur K s'il existe $\delta > 0$ tel que $\|u^* - u\| \leq \delta$ implique $J(u^*) \leq J(u)$.

Une fonction $J : H \to \mathbb{R}$ est semicontinue inférieurement (sci) si pour toute suite $\{u_n\}_{n=1}^{\infty}$ telle que $\lim u_n = u$,

$$J(u) \leq \liminf J(u_n)$$

On rappelle qu'un ensemble $K \subseteq H$ est un compact (plus précisement, K est sequentiellement compact) si à partir de toute suite $\{u_n\}_{n=1}^{\infty} \subseteq K$ on peut extraire une sous-suite $\{u_{n_k}\}_{k=1}^{\infty} \subseteq K$ telle que $\lim u_{n_k} \in K$. Un ensemble $K \subset \mathbb{R}^n$ est compact si et seulement s'il est fermé et borné. Ceci n'est plus vrai en dimension infini.

Théorème 6. *Soit J une fonction sci et K un ensemble compact. Alors J admet un minumum global sur K.*

Corollaire 3 (existence de minimum dans $\mathbb{R}^n$)**.** *Soit K un fermé de $\mathbb{R}^n$ et $J : K \to \mathbb{R}$ une fonction sci "infinie à l'infini":*

$$\lim_n J(u_n) = \infty \quad si \quad \lim_n \|u_n\| \to \infty$$

Alors J admet un minimum global sur K.

Fonctions convexes Un ensemble K est *convexe* si $\forall u, v \in K \ \forall \theta \in [0, 1]$,

$$\theta u + (1 - \theta)v \in K$$

Soit K un ensemble convexe. Une fonction $J : K \to \mathbb{R}$ est *convexe* si $\forall u, v \in K \ \forall \theta \in [0, 1]$,

$$J(\theta u + (1 - \theta)v) \leq \theta J(u) + (1 - \theta)J(v)$$

On dit que J est *strictement convexe* si l'inégalité est stricte dès que $u \neq v$ et $\theta \in (0, 1)$.

Théorème 7 (Minimisation d'une fonction convexe)**.** *Soit $K \subseteq H$ un ensemble convexe fermé et soit $J : K \to \mathbb{R}$ une fonction convexe sci "infinie à l'infini":*

$$\lim_n J(u_n) = \infty \quad si \quad \lim_n \|u_n\| \to \infty$$

Alors J admet un minimum global sur K. Si J est strictement convexe, le minimum est unique.

9.1.2 Problèmes d'optimisation en finance

Les différents problèmes d'optimisation en finance nécessitent des algorithmes différents:

1. Programmation linéaire:
$$\inf\langle a, x\rangle \quad \text{s. c.} \quad Bx \geq b$$
Exemple: optimisation de rendement de portefeuille sous contraintes de trading. Solution: méthode de simplexe.

2. Programmation quadratique:
$$\inf\langle Ax, x\rangle \quad \text{s. c.} \quad Bx \geq b$$
Exemple: minimisation de la variance de portefeuille sous contraintes de rendement et de trading.

3. Minimisation d'une fonction convexe (ou presque) et différentiable. Exemple: problèmes de calibration régularisés. Solution: méthodes de gradient, gradient conjugué ou quasi-newton.

4. Minimisation d'une fonction non convexe et/ou non différentiable. Exemple: problèmes de calibration sans régularisation. Solution possible: algorithmes évolutionnaires.

Dans cet section on se concentre sur les méthodes de minimisation des fonctions convexes.

9.1.3 Minimisation de fonctions convexes

Dans cette section, on étudie les algorithmes permettant de minimiser une fonction convexe $J : \mathbb{R}^n \to \mathbb{R}$.

- Les *méthodes de descente* consistent à faire deplacer un point dans l'espace de recherche
- Dans les algorithmes *à pas optimal*, à l'itération n, une direction de descente p^n est choisie puis on résout un problème de minimisation unidimensionnelle:
$$u^{n+1} = u^n + \lambda^* p^n, \qquad \lambda^* = \arg\inf_{\lambda\in\mathbb{R}} J(u^n + \lambda p^n)$$
- Si J n'est pas convexe, ces algorithmes trouvent un minimum local.

Un algorithme d'optimisation unidimensionnelle est donc un brique de base pour presque tous les méthodes multidimensionnelles.

Optimisation unidimensionnelle Les algorithmes de minimisation unidimensionnelle s'inspirent des méthodes de solution d'équations: bissection, méthode de Newton ou combinaison de deux. Pour démarrer, ces algorithmes nécessitent le plus souvent que le minimum soit déjà localisé: il faut trouver 3 nombres $a < b < c$ tels que $f(b) < f(a)$ et $f(b) < f(c)$. L'algorithme "golden section search", analogue à la bissection, procede ensuite par itérations successives:

- Choisir un point x, soit dans (a_n, b_n) soit dans (b_n, c_n). Supposons pour fixer les idées que $x \in (b_n, c_n)$.
- Si $f(b_n) < f(x)$, poser
$$(a_{n+1}, b_{n+1}, c_{n+1}) = (a_n, b_n, x);$$
sinon poser
$$(a_{n+1}, b_{n+1}, c_{n+1}) = (b_n, x, c_n).$$

Le fonctionnement de cet algorithme est representé graphiquement sur la fig. 9.1. La règle optimale pour le choix de x consiste à prendre

$$x = b_n + w(c_n - b_n), \quad \text{si} \quad c_n - b_n > b_n - a_n$$
$$x = b_n - w(b_n - a_n), \quad \text{sinon},$$

où

$$w = \frac{3 - \sqrt{5}}{2} \approx 0.382$$

est le 'nombre d'or'.

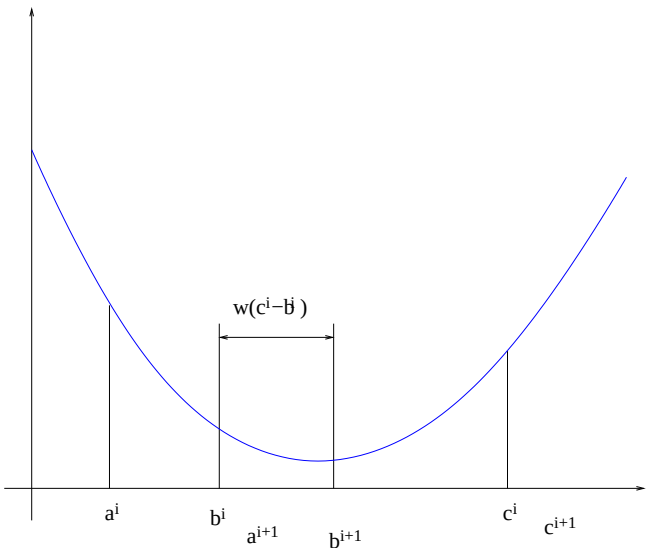


Figure 9.1: L'algorithme "golden section search" pour la recherche de minumum en une dimension.

Algorithme de relaxation La méthode la plus intuitive pour minimiser une fonction $J : \mathbb{R}^n \to \mathbb{R}$ consiste à la minimiser successivement suivant chaque dimension. Pour calculer l'itération u^{i+1} à partir de u^i, on pose,

$$\begin{aligned} u_1^{i+1} &= \arg\inf_u J(u, u_2^i, \ldots, u_n^i) \\ u_2^{i+1} &= \arg\inf_u J(u_1^{i+1}, u, u_3^i, \ldots, u_n^i) \\ &\ldots \\ u_n^{i+1} &= \arg\inf_u J(u_1^{i+1}, \ldots, u_{n-1}^{i+1}, u) \end{aligned}$$

Figure 9.2, graphique de gauche, montre le fonctionnement de cet algorithme pour la fonction

$$f(x, y) = (x + y)^2 + \frac{1}{20}(x - y)^2.$$

On voit que la performance peut être améliorée en choisissant bien la direction de descente.

Algorithme de gradient à pas optimal Dans cet algorithme, on cherche à aller toujours dans la direction de decroissance la plus rapide de la fonction J. Pour calculer l'itération u^{i+1} à partir de u^i, on pose,

$$\begin{aligned} u^{i+1} &= u^i - \lambda^i J'(u^i) \\ \text{où} \quad \lambda^i &= \arg\inf_{\lambda \in \mathbb{R}} J(u^i - \lambda J'(u^i)) \end{aligned}$$

Comme le montre fig. 9.2, graphique de droite, il n'est pas toujour judicieux de choisir la direction de descente la plus rapide!

Algorithme de gradient conjugué L'algorithme de gradient conjugué choisit la meuilleure direction en tenant compte non seulement de l'endroit où l'on se trouve, mais aussi de toutes les directions que l'on avait choisies auparavant. Supposons dans un premier temps:

$$J(u) = \frac{1}{2}(Au, u) - (b, u)$$

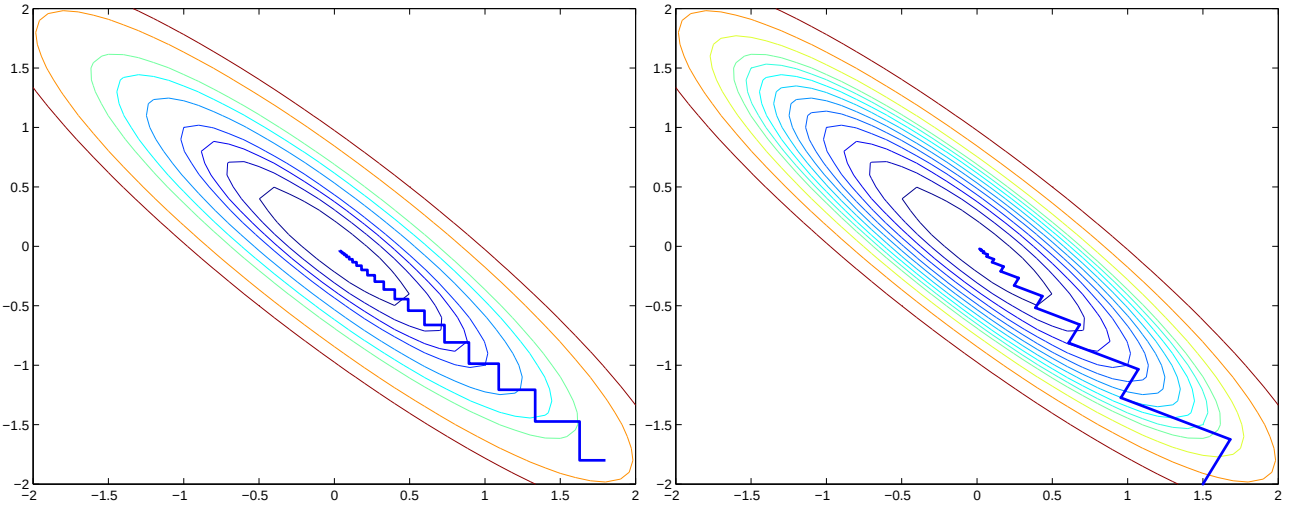


Figure 9.2: Gauche: Algorithme de relaxation (minimisation successive suivant chaque dimension). Droite: Algorithme de gradient à pas optimal. Les ellipses sur chaque graphique sont les lignes de niveau de la fonction à minimiser.

Dans ce contexte la méthode de gradient s'écrit:

$$u^{n+1} = \arg \inf_{u \in u^n + \mathcal{L}(J'(u^n))} J(u)$$

La méthode de *gradient conjugué* propose:

$$u^{n+1} = \arg \inf_{u \in u^n + \mathcal{L}(J'(u^0), \dots, J'(u^n))} J(u)$$

La condition de 1er ordre implique:

$$\forall n, \quad \langle J'(u^n), J'(u^k) \rangle = 0 \quad \forall k < n$$

Soit $p^n := u^{n+1} - u^n$ le n-ième pas de l'algorithme. Alors

$$\langle Ap^n, J'(u^k) \rangle = 0 \quad \forall k < n$$

Ceci implique que $Ap^n \in \mathcal{L}(J'(u^0), \dots, J'(u^{n-1}))^\perp$ et donc

$$\langle Ap^n, p^k \rangle = 0 \quad \forall k < n \tag{9.1}$$

Les vecteurs $(p^k)_{1 \le k \le n}$ sont dits conjugués par rapport à A.

Pour chaque n, il existe $x \in \mathcal{L}(J'(u^0), \dots, J'(u^{n-1}))$ tel que

$$p^n = \alpha J'(u^n) + x$$

La relation de conjugaison (9.1) implique que pour tout $k < n-1$, $\langle Ax, p^k \rangle = 0$ et donc $x \sim p^{n-1}$. Il existent alors z^n, $k^n \in \mathbb{R}$ et $\lambda^n \in \mathbb{R}$ avec

$$p^n = \lambda^n z^n \quad \text{et} \quad z^n = J'(u^n) + k^n z^{n-1}$$

La relation (9.1) donne à nouveau: $\langle z^n, Az^{n-1} \rangle = 0$ et

$$k^n = -\frac{\langle J'(u^n), Az^{n-1} \rangle}{\langle z^{n-1}, Az^{n-1} \rangle} = \frac{\|J'(u^n)\|^2}{\|J'(u^{n-1})\|^2}$$

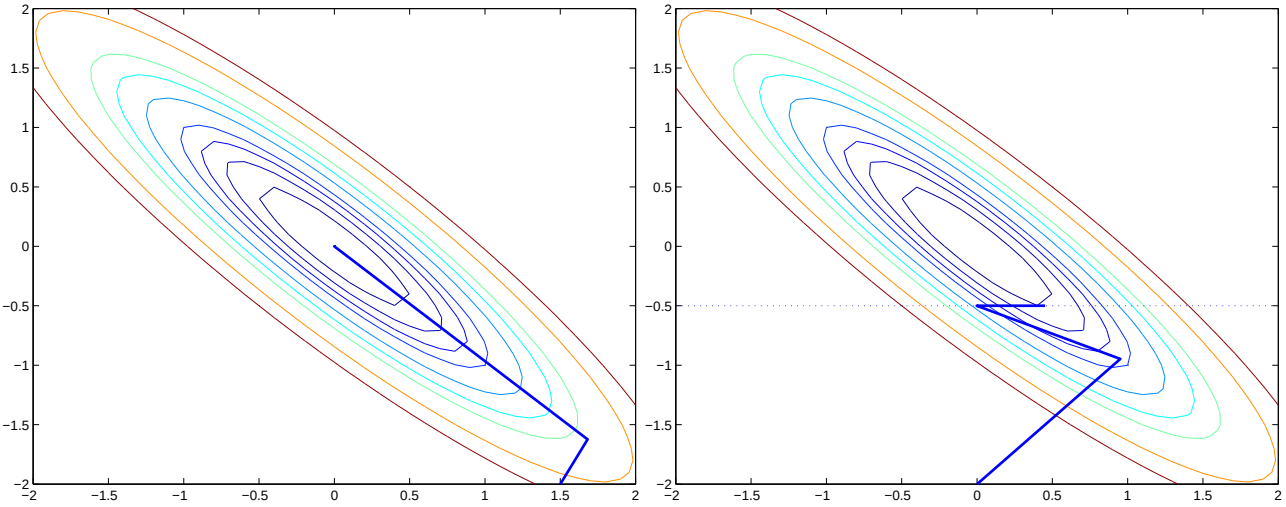


Figure 9.3: Gauche: méthode de gradient conjugué. Droite: méthode de gradient avec projection.

Version finale de l'algorithme:
Initialisation: u^0, $z^0 = J'(u^0)$

$$\begin{aligned} u^{n+1} &= u^n + \lambda z^n \\ \text{où} \quad \lambda &= \arg \inf_{\lambda \in \mathbb{R}} J(u^n + \lambda z^n) \\ z^{n+1} &= J'(u^{n+1}) + \frac{\|J'(u^{n+1})\|^2}{\|J'(u^n)\|^2} z^n \end{aligned}$$

L'algorithme permet de minimiser une forme quadratique N-dimensionnelle en N itérations (voir fig. 9.3, graphique de gauche). Il est aussi efficace pour des fonctions convexes générales.

Méthodes de type Newton Dans les méthodes de type Newton, on utilise non seulement le gradient de J mais aussi sa matrice Hessienne J''. Pour calculer l'itération u^{i+1} à partir de u^i, on écrit un développement de J autour de u^i:

$$J(u) \approx J(u^i) + J'(u^i)(u - u^i) + \frac{1}{2}(u - u^i)^* J''(u^i)(u - u^i)$$

Cette expression est minimale pour

$$u = u^i - \{J''(u^i)\}^{-1} J'(u^i)$$

La méthode de Newton consiste à poser

$$u^{i+1} = u^i - \{J''(u^i)\}^{-1} J'(u^i)$$

Avantages des méthodes de type Newton

- Minimisation d'une forme quadratique en une itération.
- Convergence quadratique près de minimum.

Les défauts

- Le calcul et l'inversion de J'' est nécessaire.
- La performance est mauvaise si J'' n'est pas définie positive ou si J ne ressemble pas à une forme quadratique.

Solution: méthodes *quasi-Newton* (e.g. BFGS): à chaque itération, on utilise J' pour mettre à jour une approximation de J'' qui reste toujours définie positive.

Optimisation sous contraintes On cherche à résoudre

$$\inf_{u\in K} J(u),$$

où K est typiquement un ensemble convexe fermé. En cas de contraintes d'égalite, on essaie, si possible de passer aux variables indépendantes ou introduire les multiplicateurs de Lagrange. En cas de contraintes d'inégalité: méthodes de projection ou pénalisation de contraintes.

Méthode de gradient avec projection L'algorithme de gradient à pas optimal avec projection consiste à projetter la solution sur l'ensemble admissible après chaque itération de la méthode de gradient. Pour calculer l'itération u^{i+1} à partir de u^i, on pose,

$$u^{i+1} = P_K(u^i - \lambda^i J'(u^i)),$$
$$\text{où} \quad \lambda^i = \arg\inf_{\lambda\in\mathbb{R}} J(u^i - \lambda J'(u^i)).$$

Si l'ensemble K est un rectangle

$$K = [a_1, b_1]\dots[a_n, b_n],$$

l'opérateur de projection a la forme particulièrement simple:

$$[P_K u]_i = \min(b_i, \max(a_i, u_i)).$$

Figure 9.3, graphique de droite, illustre le fonctionnement de cet algorithme pour l'ensemble de contraintes $K = \{(x, y) : y \le -0.5\}$.

Méthodes de pénalisation de contraintes Soit K de la forme

$$K = \{u : F_i(u) \le 0, i = 1, \dots M\}$$

avec les fonctions F_i convexes et continues. La méthode de pénalisation de contraintes consiste à remplacer le problème initial par un nouveau problème sans contraintes et avec une fonction d'objectif modifiée:

$$u_\varepsilon = \arg\inf_{v\in\mathbb{R}^n} \left(J(v) + \frac{1}{\varepsilon}\sum_{i=1}^{M} \max(F_i(v), 0) \right).$$

Le problème penalisé reste alors convexe et on peut démontrer que

$$\lim_{\varepsilon\to 0} u_\varepsilon = \arg\inf_{v\in K} J(v).$$

Bibliographie et liens .

- Méthodes numériques générales; description et code source de nombreux algorithmes d'optimisation sans contrainte
  W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: the Art of Scientific Computing*. Livre disponible sur internet: `www.nr.com`
- Implémentation professionnelle de la méthode BFGS avec contraintes de type intervalle
  C. Zhu, R. H. Byrd and J. Nocedal, *L-BFGS-B: Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimisation.*
  `www.ece.northwestern.edu/~nocedal/lbfgsb.html`

Appendix A

Quelques rappels sur le calcul stochastique et les processus de Markov

Théorème 8 (Critère de Novikov)**.** *Soit M une martingale locale continue et supposons que*

$$E[e^{\frac{1}{2}\langle M\rangle_\infty}] < \infty$$

Alors $\mathcal{E}(M) := (e^{M_t - \frac{1}{2}\langle M\rangle_t})_{t\geq 0}$ est une martingale uniformement intégrable.

Approximation d'un processus de diffusion par une chaîne de Markov Soit X une diffusion markovienne:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t,$$

où W est un mouvement standard, et pour tout $h > 0$, soit $\{X_n^h\}_{n\geq 1}$ une chaîne de Markov. On dit que la famille de chaînes de Markov $\{X_n^h\}_{n\geq 1}^{h>0}$ est localement consistante avec X si elle satisfait les conditions suivantes:

$$\begin{aligned}
&E[\Delta X_n^h | X_n^h] = \mu(X_n^h)h + o(h),\\
&\text{Var}\,[\Delta X_n^h | X_n^h] = a(X_n^h)h + o(h),\\
&\sup_{n,\omega} |\Delta X_n^h| \overset{h\to 0}{\longrightarrow} 0,
\end{aligned}$$

où on utilise la notation $\Delta X_n^h := X_{n+1}^h - X_n^h$ et $a(x) = \sigma(x)\sigma(x)^\perp$.

Théorème 9 (Approximation par chaîne de Markov, [43])**.** *Soit σ et μ continus et bornés et soit $\{X_n^h\}_{n\geq 1}^{h>0}$ une famille de chaînes de Markov faiblement consistante avec X. Alors le processus interpolé défini par*

$$\tilde{X}_t^h = X_{[t/h]}^h$$

converge faiblement vers une diffusion markovienne avec dérive μ et coefficient de diffusion σ.

Appendix B

Examens des années précedentes

B.1 Master Modélisation Aléatoire, examen du 27 avril 2006

Exercice 1 (Problèmes inverses mal posés; 6 points)**.** On cherche à résoudre l'équation

$$AX = E \tag{B.1}$$

avec

$$A = \begin{pmatrix} \varepsilon & 0 & 0 \\ \varepsilon & 1 & 1 \\ \varepsilon & -1 & -1 \end{pmatrix} \qquad E = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{et} \quad X \in \mathbb{R}^3$$

1. Cette équation, a-t-elle une solution au sens strict? Calculer l'ensemble de ces solutions au sens de moindres carrés. Calculer la solution moindres carrés de norme minimale.

2. Quelle sera la solution moindres carrés de norme minimale si on remplace E par

$$E^\delta = \begin{pmatrix} 1+\delta \\ 1 \\ 0 \end{pmatrix}$$

   Que peut-on dire sur la stabilité de calcul lorsque $\varepsilon \to 0$.

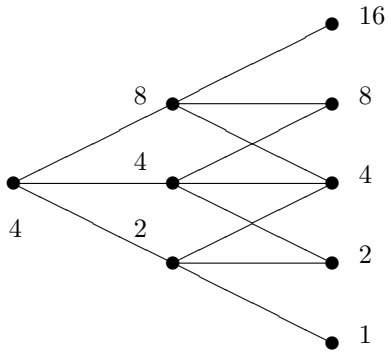
3. Calculer la solution de ce problème pour la donnée E^δ en utilisant la régularisation de Tikhonov avec paramètre de régularisation α donnée par

$$X_\alpha^\delta = \arg\inf_X \{\|AX - E^\delta\|^2 + \alpha\|X\|^2\}$$

   pour un $\alpha > 0$ fixé, que peut-on dire sur la stabilité de calcul lorsque $\varepsilon \to 0$.

4. Supposons que $\alpha(\delta)$ est choisi tel que $\alpha(\delta) \to 0$ quand $\delta \to 0$. Montrer que $X_{\alpha(\delta)}^\delta \to X^\dagger$ lorsque $\delta \to 0$, où $X^\dagger$ est la solution de norme minimale de (B.1). Pourquoi, dans ce cas on n'a pas besoin d'imposer une condition supplémentaire (voir le cours) sur $\alpha(\delta)$ pour avoir la convergence?

Exercice 2 (Arbres trinomiaux implicites; 5 points)**.** Dans cet exercice on considère un arbre trinomial à deux périodes de structure suivante (les prix de l'action S dans les noeuds sont indiqués sur l'arbre):



Sur ce marché, les prix d'options suivants ont été observés:

- Un call de strike $K = 4$ et maturité $T = 1$ cote à 1.2.
- Un call de $K = 2$ et $T = 2$ cote à 2.36.
- Un call de $K = 4$ et $T = 2$ cote à 1.3.
- Un call de $K = 8$ et $T = 2$ cote à 0.48.

1. En supposant que le taux d'intérêt est égal à zéro et en utilisant la méthode décrite dans le cours, calculer les probabilités de transition dans cet arbre (12 probabilités au total).

2. Utiliser ces probabilités de transition pour calculer le prix d'une option Call up–and–out avec une barrière à $B = 6$ (la valeur de l'option est zéro à maturité si la barrière a été franchi pendant la vie de l'option), de strike $K = 2$ et maturité $T = 2$.

Problème 1 (9 points)**.** Dans un marché financier soumis au défaut il existent deux actifs risqués:

- Une action S dont l'évolution du prix est décrit par

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t, \quad t < \tau, \qquad S_t = 0, \quad t \geq \tau, \tag{B.2}$$

  où τ est l'instant du défaut.

- Un zéro-coupon B soumis au défaut de maturité T. A la maturité, le prix de ce zéro-coupon vaut 1 si $\tau > T$ et zéro sinon. A n'importe quelle date le prix de ce zéro-coupon est une fonction uniquement du temps et de la valeur de l'action notée par $B(t, S)$.

Il existe également un compte bancaire sans risque dont le taux d'intérêt est constant et égal à r.

Première partie Dans un premier temps on cherche à démontrer qu'une option européenne sur S peut être couvert par une portefeuille contenant des actions, des zéro-coupons risqués et l'actif sans risque.

1. Soit V_t la valeur du portefeuille autofinançant contenant du cash, δ_t actions et ω_t unités de B. Ecrire l'expression pour dV_t.

2. On considère maintenant une option sur S qui paie $h(S)$ à la date T (maturité). Soit $C(t, S)$ le prix de cet option à une date $t < T$. On cherche à construire un portefeuille autofinançant dont la valeur réplique $C(t, S_t)$. En utilisant le fait que $dC(t, S_t) = dV_t$ avant le défaut et que $\Delta C(t, S_t) = \Delta V_t$ à l'instant du défaut, déterminer δ_t, ω_t et l'équation satisfaite par $C(t, S)$.

3. Montrer que l'équation pour $C(t, S)$ peut s'écrire sous la forme

$$\frac{LC - rC + rS\frac{\partial C}{\partial S}}{C(t,S) - C(t,0) - S\frac{\partial C}{\partial S}} = \frac{LB - rB + rS\frac{\partial B}{\partial S}}{B(t,S) - S\frac{\partial B}{\partial S}} := \lambda(t, S), \tag{B.3}$$

   où $LC = \frac{\partial C}{\partial t} + \frac{1}{2}S^2\sigma^2\frac{\partial^2 C}{\partial S^2}$.

4. Rappeler l'équation de Black-Scholes. Montrer que l'équation (B.3) est une équation de Black-Scholes avec un terme source et avec un taux d'intérêt modifié, à définir.

Deuxième partie Dans cette partie on suppose que $\lambda(t,S) = \lambda$ **est une constante.**

1. Montrer que par un changement de variable on peut ramener l'équation (B.3) à une équation de Black-Scholes sans terme source, c'est-à-dire, montrer qu'on peut toujours trouver une décomposition $C(t,S) = \tilde{C}(t,S) + f(t)$ telle que $\tilde{C}$ satisfait l'équation de Black-Scholes sans terme source, où f est une fonction à définir.

2. En utilisant le résultat précédent, donner une formule explicite pour la solution de (B.3) dans le cas des options call et put européennes.

3. Calculer également le prix $B(t,S)$ du zéro-coupon soumis au défaut. En déduire que λ est toujours positif.

4. Montrer que dans ce modèle en présence de possibilité de défaut, les options call sont toujours plus chers que dans le modèle Black-Scholes standard.

5. Montrer que le prix d'une option européenne de payoff $h(S_T)$ peut s'écrire comme l'espérance actualisée de $h(S_T)$ sous la probabilité Q sous laquelle

$$\frac{dS_t}{S_t} = (r+\lambda)dt + \sigma dW_t^Q, \quad t < \tau, \qquad S_t = 0, \quad t \geq \tau,$$

où W^Q est un Q-mouvement Brownien et τ est une variable exponentielle indépendante de W de paramètre λ.

B.2 Master Modélisation Aléatoire, examen du 10 mai 2007

Exercice 1 (Arbre trinomial, 6 points)**.** Dans un marché incomplet sans arbitrage, le prix de certains actifs n'est pas unique, mais on peut toujours calculer les bornes d'arbitrage sur leur prix, données par

$$\inf_{Q\in\mathcal{Q}} E^Q[X] \leq \pi(X) \leq \sup_{Q\in\mathcal{Q}} E^Q[X],$$

où $\mathcal{Q}$ est l'ensemble de toutes les probabilités risque-neutres sur le marché, X la valeur terminale de l'actif, $\pi(X)$ son prix à l'instant initial, et le taux d'intérêt est supposé nul. Dans cet exercice on calculera les bornes d'arbitrage pour le prix d'une option dans un marché trinomial.

On se place dans un premier temps dans un modèle trinomial à 1 période: un marché contenant un actif sans risque B avec $B_0 \equiv B_1$ et un actif risqué S où S_1 peut prendre les valeurs $\{dS_0, S_0, uS_0\}$, $d < 1 < u$. On note f la fonction pay-off de l'option qui est supposée convexe comme c'est le cas pour les calls et les puts.

1. Décrire l'ensemble des probabilités risque-neutres dans ce marché; montrer que cet ensemble peut être paramétré par la probabilité p de transition $S_0 \to uS_0$.

2. Montrer que l'espérance $E[f(S_1)]$ est une fonction croissante de p. En déduire les bornes d'arbitrage $\inf_{Q\in\mathcal{Q}} E^Q[f(S_1)]$ et $\sup_{Q\in\mathcal{Q}} E^Q[f(S_1)]$.

3. Montrer que les deux bornes sont des fonctions convexes de S_0 et que chaque borne peut s'écrire comme le prix d'une option de pay-off f dans un modèle plus simple qui constitue un cas particulier dégénéré du modèle trinomial.

On se place maintenant dans un modèle trinomial à T périodes.

4. Calculer les bornes d'arbitrage $\inf_{Q\in\mathcal{Q}} E^Q[f(S_T)]$ et $\sup_{Q\in\mathcal{Q}} E^Q[f(S_T)]$. Pour la borne supérieure, la solution ne sera pas complètement explicite, il s'agit de donner la méthode de calcul.

   Indication: utiliser la convexité et les espérances conditionnelles emboîtées pour se ramener au cas traité dans 1.

Exercice 2 (Options forward start, 5 points)**.** Les options forward start sont les options dont le strike est déterminé à une date future selon une règle spécifique, par exemple,

$$H_T = (S_T - mS_{T_0})^+, \tag{1}$$

où $T_0 < T$ est une date future et m est un nombre fixé dans le contrat (moneyness de l'option). On se place dans un modèle exponentielle Lévy sous la probabilité risque-neutre:

$$S_t = S_0 e^{rt + X_t},$$

où (X_t) est un processus de Lévy, et on suppose que les prix des options Européennes de tous les strikes et toutes les maturités sont observables sur le marché.

1. Calculer (exprimer à partir des prix des européens) le prix à l'instant $t = 0$ d'une option forward start avec pay-off donné par (1).
   Indication: utiliser l'indépendance et la stationnarité des accroissements d'un processus de Lévy.

2. Une option cliquet (séquence d'options forward start) a un pay-off de

$$\sum_{i=1}^{12} \max(0, r_i), \quad r_i = S_{\frac{i}{12}} - S_{\frac{i-1}{12}}.$$

   Calculer le prix à l'instant $t = 0$ de cette option. Est-ce que le résultat vous paraît réaliste? Conclure sur la pertinence des modèles exponentielle-Lévy pour le pricing des options cliquets.

Exercice 3 (Couverture dans un modèle exp-Lévy, 6 points)**.** On se place dans un modèle exponentielle Lévy sous la probabilité martingale et en l'absence de taux d'intérêt:

$$S_t = S_0 e^{X_t},$$

où (X_t) est un processus de Lévy de la forme

$$X_t = \gamma t + \sigma W_t + aN_t,$$

avec W un mouvement Brownien et N un processus de Poisson standard d'intensité λ.

1. Calculer la valeur de γ pour laquelle (e^{X_t}) devient une martingale.

On souhaite démontrer que dans ce modèle on peut couvrir parfaitement une option européenne de pay-off H_T avec une stratégie dynamique de trading en actions et une autre option, de pay-off H_T^0. On note par $C^0(t, S_t) = E[H_T^0 | \mathcal{F}_t]$ le prix de l'option de couverture et par $C(t, S_t) = E[H_T | \mathcal{F}_t]$ le prix de l'option qu'on cherche à couvrir.

2 Ecrire la dynamique de $C(t, S_t)$ et $C^0(t, S_t)$ et celle d'un portefeuille autofinançant contenant δ_t actions et ω_t unités de l'option C^0. La valeur initiale du portefeuille est égale à $C(0, S_0) = E[H_T]$.

3 En utilisant la relation d'isométrie pour les intégrales stochastiques, calculer l'espérance du carré de l'erreur de couverture (différence entre le portefeuille de couverture et H_T). Montrer qu'il est possible de choisir des processus (δ_t) et (ω_t) pour annuler l'erreur de couverture. De quelle condition sur C^0 avez-vous besoin?

4 Montrer que lorsque la taille du saut a est petite, les ratios de couverture sont proches de ratios delta-gamma obtenus dans le modèle Black-Scholes.

Exercice 4 (Swap de volatilité, 7 points)**.** On se place dans un modèle à volatilité stochastique sous la probabilité risque-neutre et avec taux d'intérêt nul. La dynamique du sous-jacent s'écrit

$$\frac{dS_t}{S_t} = \sigma_t dW_t.$$

En introduisant le processus du log-prix $X_t = \log(S_t/S_0)$, on a

$$\langle X \rangle_t = \int_0^t \sigma_s^2 ds.$$

Le *swap de volatilité* est un contrat d'échange de la volatilité réalisée d'un actif sur une période contre un somme fixe. Dans l'approximation continue, le pay-off d'un swap de volatilité s'écrit

$$\left(\frac{1}{T} \int_0^T \sigma_s^2 ds \right)^{1/2} - \sigma_{vol},$$

où σ_{vol} est une valeur spécifiée dans le contrat. Il est rappelé que le pay-off du swap de variance, traité en cours, est donné par

$$\frac{1}{T} \int_0^T \sigma_s^2 ds - \sigma_{var}^2,$$

Calculer le taux de swap de volatilité signifie trouver σ_{vol} qui annule la valeur du swap à l'instant $t = 0$.

Vous pouvez imposer les hypothèses d'intégrabilité que vous trouvez nécessaires, à condition qu'elles soient clairement identifiées dans votre solution.

1. Montrer que le pay-off d'un swap de volatilité est majoré par le pay-off du swap de variance sur le même sous-jacent plus une constante. En utilisant ce résultat, donner une borne supérieure pour le taux du swap de volatilité σ_{vol} en fonction du taux du swap de variance σ_{var} et le portefeuille de surreplication correspondant.

Dans la suite de l'exercice on suppose de plus que la volatilité (σ_t) est indépendante du mouvement brownien (W_t).

2. Ecrire la dynamique de X et montrer que, sous l'hypothèse d'indépendance ci-dessus,

$$E[e^{\lambda \langle X \rangle_t}] = E[e^{\left(\frac{1}{2} - \sqrt{\frac{1}{4} + 2\lambda}\right) X_t}], \quad \forall \lambda \in \mathbb{C}.$$

3. On rappelle la représentation suivante pour la fonction racine:

$$\sqrt{q} = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - e^{-sq}}{s^{3/2}} ds, \quad q \geq 0.$$

   Utiliser cette représentation pour montrer que $E[\langle X \rangle_T^{1/2}] = E[h(S_T/S_0)]$ avec h une fonction déterministe donnée par

$$h(x) = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1 - \mathrm{Re}[e^{\left(\frac{1}{2} - \sqrt{\frac{1}{4} - 2s}\right) \log x}]}{s^{3/2}} ds.$$

   En déduire une méthode pour calculer le taux de swap de volatilité à partir de la densité risque-neutre $q(\cdot)$.

Pour la suite de l'exercice on se place dans le cadre du modèle de Heston vu en cours, avec paramètres $k = 1$, $\theta = 0.01$, $\delta = 0.1$ $\rho = 0$ et $v_0 = 0.01$.

4. Calculer le taux de swap de variance d'échéance 3 mois, et la borne supérieure pour le taux de swap de volatilité (que vous avez obtenu en question 1 ci-dessus).

5. Calculer une approximation pour le taux de swap de volatilité, en utilisant une "correction de convexité": décomposition de la fonction $\sqrt{x}$ en série de Taylor jusqu'à l'ordre 2 autour du point $E[\langle X \rangle_T]$.
   Indication pour le calcul de variance de $\langle X \rangle_T$ vous pouvez supposer que T est petit.

B.3 Master Modélisation Aléatoire, examen du 14 mai 2008

Exercice 1 (Corridor variance swaps)**.** Un corridor variance swap est une modification du variance swap où la variance du sous-jacent est comptabilisé pour le calcul du pay-off seulement lorsque le sous-jacent se trouve entre les bornes K_1 et K_2. Le pay-off (dans l'approximation continue) est ainsi donné par

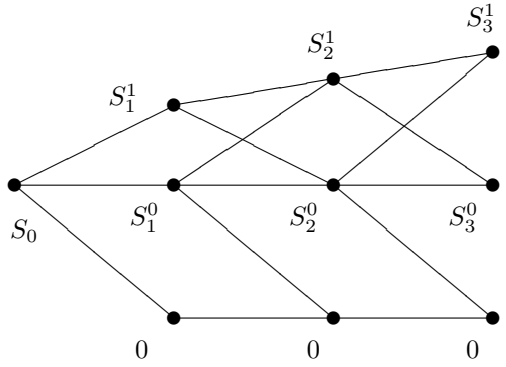
$$H_T = \frac{1}{T}\int_0^T \sigma_t^2 1_{[K_1,K_2]}(S_t)dt - \sigma_K^2,$$

où σ_K est la valeur contractuelle de volatilité. Dans cet exercice on suppose que le processus de prix S est continu et que **le taux d'intérêt est nul**.

1. En appliquant la formule d'Itô à $f(S_t)$, montrer qu'il est possible, en choisissant une fonction f adaptée, de représenter le pay-off d'un corridor variance swap comme la somme d'une fonction de S_T et d'une intégrale stochastique par rapport à S_t.

2. En déduire une méthode de réplication du corridor variance swap avec un portefeuille statique d'options européennes et une stratégie dynamique de trading en l'actif sous-jacent. Calculer également σ_K qui annule la valeur du swap.

Exercice 2 (Arbres implicites)**.** On considère un modèle discret à trois périodes avec possibilité de défaut, dont la structure est représentée sur le graphique suivant. Les numéros des noeuds sont indiqués sur l'arbre. Les prix du sous-jacent dans les noeuds sont $S_0 = S_1^0 = S_2^0 = S_3^0 = 80$, $S_1^1 = 100$, $S_2^1 = 120$ et $S_3^1 = 140$.



On supposera le taux d'intérêt nul dans tout l'exercice. Dans la première partie, on considère que le processus S est markovien, c'est-à-dire, les probabilités de transition dépendent seulement du noeud, mais pas de la trajectoire empruntée pour arriver à ce noeud (comme dans le cours).

Les probabilités de transition entre les différents noeuds sont notées p_n^m (vers le haut) et q_n^m (vers le bas) avec n et m les indices du noeud correspondant.

1. A partir de la condition de non-arbitrage (S martingale), déterminer les probabilités p_1^1, p_1^2, q_1^1, q_1^2 et les contraintes sur les probabilités p_0, q_0, p_1^0, q_1^0, p_2^0, q_2^0. Combien reste-t-il de degrés de liberté dans ce modèle?

2. Les options suivantes sont cotées à la date $T = 0$:

   - Un Call à la monnaie d'échéance $T = 1$, au prix de 12.
   - Un Call à la monnaie d'échéance $T = 2$, au prix de 15.
   - Un Call à la monnaie d'échéance $T = 3$, au prix de 20.25.

   Déterminer les probabilités p_0, q_0, p_1^1, q_1^1, p_2^1, q_2^1.

3. Utiliser les probabilités de transition pour calculer le prix d'une option Put de strike $K = 90$ d'échéance $T = 3$ avec une barrière désactivante à 100.

Maintenant, on ne supposera plus le processus S markovien, c'est-à-dire, les probabilités de transition à partir d'un noeud peuvent dépendre du chemin que le processus de prix a emprunté pour arriver à ce noeud.

4. Montrer, à partir de la condition d'absence d'arbitrage, que pour tous les noeuds sauf S_2^0, les probabilités de transition sont indépendantes de trajectoire et ont la même valeur que dans la 1ère partie de l'exercice.

5. On notera par $\hat{p}_2^0$ et $\bar{p}_2^0$ la valeur de la probabilité p_2^0 dans le cas ou le processus de prix a emprunté, respectivement, une trajectoire haussière ou plate pour arriver au noeud S_2^0. En utilisant la connaissance du prix du Call à la monnaie d'échéance $T = 3$, démontrer que $24\hat{p}_2^0 + 11\bar{p}_2^0 = 7$. Combien de degrés de liberté reste-t-il dans le modèle?

6. Exprimer le prix d'une option Put de strike $K = 90$ d'échéance $T = 3$ avec une barrière désactivante à 100 en fonction de $\hat{p}_2^0$ et $\bar{p}_2^0$. En déduire les bornes de non-arbitrage pour le prix de cette option.

Exercice 3 (Crash notes)**.** Une "crash note" est une option exotique permettant de se couvrir contre le risque d'une forte chute quasi-instantanée du sous-jacent. Dans cet exercice on considère une variante de ce produit appelée "daily cliquet put spread". Les caractéristiques de cette option sont l'échéance T, les deux niveaux de strike $K_1 < K_2 < 1$ et la valeur nominale N. On note par R_t le rendement instantané[1] du sous-jacent $R_t = \frac{S_t}{S_{t-}}$ et par τ la première date a laquelle ce rendement est inférieur à K_2:

$$\tau = \inf\{t : R_t \leq K_2\}.$$

Le pay-off du produit est défini comme suit:

- Si $R_t > K_2$ pour tout $t \leq T$, le pay-off est nul.
- Sinon, l'acheteur reçoit à la date τ le montant $N(K_2 - R_\tau)^+ - N(K_1 - R\tau)^+$ et le produit s'arrête.

Pour l'étude de ce produit, on se dans le cadre d'un modèle exponentielle-Lévy **risque-neutre**: $S_t = S_0 e^{X_t}$, où X est un processus de Lévy de triplet caractéristique (σ^2, ν, γ) tel que S_t est une martingale. Le taux d'intérêt est supposé nul.

1. Exprimer le pay-off du produit en fonction des variables τ et ΔX_τ.

2. Montrer que τ et ΔX_τ sont indépendants, et donner les lois de ces deux variables.

3. En déduire le prix du daily cliquet (en fonction de la mesure de Lévy ν de X). Quel est le delta de ce produit?

4. Calculer le prix du daily cliquet dans le modèle de Kou, où la mesure de Lévy a une densité

$$\nu(x) = c_+ e^{-x/\eta_+} 1_{x>0} + c_- e^{x/\eta_-} 1_{x<0}.$$

5. Donner la représentation du pay-off du daily cliquet comme l'intégrale d'une fonction à préciser par rapport à la mesure aléatoire de Poisson de X, $J_X(ds \times dz)$.

6. En déduire une représentation de martingale pour le prix C_t du daily cliquet en fonction de temps t (autrement dit, il faut représenter le prix comme une constante + intégrale par rapport à la mesure aléatoire compensée de X, $\tilde{J}_X(ds \times dz)$).

7. En déduire le ratio de couverture optimal quadratique pour le daily cliquet. En comparant ce ratio au delta du produit, conclure sur l'applicabilité de la couverture en delta en présence de sauts.

[1] En réalité on prend plutôt le rendement journalier, d'où le nom daily cliquet, mais dans ce problème on utilise le rendement instantané pour simplifier le traitement.

B.4 Master Modélisation Aléatoire, examen du 7 mai 2009

Exercice 1 (Options sur variance)**.** Dans cet exercice, S_t est un processus de prix d'action de dynamique

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t,$$

où σ est un processus stochastique. La variance intégrée de $X_t := \log S_t$ est notée par

$$V_t = \int_0^t \sigma_s^2 ds,$$

et τ_Q est le 1er instant où la variance atteint une valeur donnée : $\tau_Q = \inf\{t : V_t = Q\}$.

Dans tout l'exercice le taux d'intérêt est supposé nul.

Première partie. Dans cette partie de l'exercice on cherche à calculer le prix et la stratégie de réplication du "Perpetual timer option" dont le pay-off à l'instant τ_Q est donné par $f(S_{\tau_Q})$ avec f une fonction déterministe.

1. Soit $h(t, y)$ une fonction $C^{1,2}$. Appliquez la formule d'Itô à $h(V_t, S_t)$.

2. Montrer que si h vérifie l'EDP de Black-Scholes avec volatilité constante égale à 1, $h(Q, S_{\tau_Q})$ s'écrit comme une intégrale stochastique par rapport à dS.

3. Donner une condition terminale que h doit vérifier pour que cette intégrale puisse s'interpréter comme un portefeuille de réplication du "perpetual timer option".

4. Donner une représentation de h comme intégrale par rapport à la densité gaussienne.

5. En déduire la stratégie de réplication et la formule de prix d'un "perpetual timer option".

Deuxième partie. Maintenant on cherche une méthode de sous-réplication pour une option call sur variance réalisée, de pay-off $(V_T - Q)^+$ à l'instant T.

6. Soit λ une fonction convexe C^2 vérifiant $\lambda'' \leq 2/y^2$. En appliquant la formule d'Itô à $\lambda(S_T)$, donnez une stratégie de sous-réplication de la variance V_T avec une option européenne dont le pay-off est à préciser et un portefeuille dynamique d'actions.

7. Montrer que si $\tau_Q \leq T$ alors à l'instant τ_Q le problème de sous-réplication du payoff $(V_T - Q)^+$ devient équivalent au problème de sous-réplication d'un swap de variance (de strike nul). En utilisant la question 6., donner une stratégie de sous-réplication pour ce swap de variance avec une option européenne et un portefeille dynamique d'actions (donner la composition du portefeuille à partir de l'instant τ_Q et le prix du portefeuille à l'instant τ_Q).

8. Toujours dans le cas $\tau_Q \leq T$, en utilisant la question 5., donner une stratégie de sous-réplication pour le call sur variance réalisée sur $[0, T]$ et le prix de cette stratégie en 0 (toujours avec une option européenne et un portefeuille dynamique d'actions).

9. Montrer que cette stratégie reste une stratégie de sous-réplication même si $\tau_Q > T$.

10. En utilisant l'intégration par parties, réécrire la borne inférieure pour le prix d'une option sur variance réalisée comme une intégrale de la forme $\int_0^\infty \lambda''(y)\{\ldots\}dy$.

11. En déduire la fonction $\bar{\lambda}$ qui maximise la borne inférieure (sous les contraintes de convexité et $\lambda''(y) \leq 2/y^2$).

Exercice 2 (Valorisation d'options de courte maturité)**.** On se place dans le cadre d'un modèle exponentielle-Lévy

$$S_t = e^{X_t},$$

où X est un processus de Lévy d'intensité finie de sauts, c'est-à-dire de la forme

$$X_t = bt + \sigma W_t + \sum_{k=1}^{N_t} Y_i,$$

sous la probabilité risque-neutre Q, où N est un processus de Poisson standard et (Y_i) sont des variables i.i.d. La mesure de Lévy de X est notée par ν. Le taux d'intérêt est supposé nul. On étudie le comportement du prix d'une option Put européenne sur S de strike $K < 1$ (hors la monnaie) et de maturité T, donné par

$$P(T, K) = E[(K - S_T)^+],$$

lorsque $T \to 0$.

Les quatre questions sont indépendantes

1. Calculer la valeur de b pour laquelle S est une martingale.

2. Démontrer que
$$\lim_{t\to 0} \frac{1}{t} E[(K - e^{bt+\sigma W_t})^+] = 0.$$
Vous pouvez par exemple appliquer la règle de l'Hôpital.

3. Montrer que
$$\lim_{t\to 0} E[(K - e^{bt+\sigma W_t+Y_1})^+] = E[(K - e^{Y_1})^+].$$

4. En conditionnant l'espérance par rapport à N_t et en utilisant les deux résultats ci-dessus, démontrer que
$$\lim_{T\to 0} \frac{1}{T} P(T, K) = \int_{\mathbb{R}} (K - e^x)^+ \nu(dx).$$

B.5 Master Probabilités et Finance, examen du 27 mars 2009

Modèle à volatilité stochastique de Barndorff-Nielsen and Shephard

Première question On considère la dynamique suivante pour la volatilité:

$$d\sigma_t^2 = -k\sigma_t^2 dt + dZ_t, \tag{B.4}$$

où Z est un processus de Poisson composé d'intensité λ, dont les sauts ont une distribution exponentielle de densité exponentielle $f(z) = \eta e^{-\eta z} 1_{z>0}$. On note $v_t = \sigma_t^2$.

1. En appliquant la formule d'Itô à $e^{kt} v_t$, résoudre l'équation (B.4).

2. Calculer la fonction caractéristique $\phi_t(u) = E[e^{iuZ_t}]$ et la transformée de Laplace $L_t(u) = E[e^{-uZ_t}]$ de Z_t.

3. Calculer
$$E\left[e^{-\int_0^T f(s)dZ_s}\right]$$
pour une fonction continue bornée positive f. Indication: commencer par considerer une fonction f constante par morceaux puis passer à la limite.

4. En déduire la transformée de Laplace de σ_t^2 et $I_T := \int_0^T \sigma_s^2 ds$.

Deuxième question On considère maintenant le modèle du comportement d'un actif financier

$$dS_t = S_t \sigma_t dW_t, \tag{B.5}$$

où σ est donnée par l'équation (B.4) et W est un mouvement brownien standard sous la probabilité risque-neutre Q, indépendant de Z. Le taux d'intérêt est nul.

1. En utilisant la formule d'Itô, déduire l'EDS satisfaite par $X_t := \log S_t$.
2. Montrer que X_T a la même loi qu'un mouvement Brownien avec drift, changé de temps par un processus à préciser. En utilisant l'indépendance entre Z et W et le point 4 de la première question, calculer la fonction caractéristique de X_T.
3. En déduire les indices d'explosion de moments (pour T fixé)

$$\tilde{p} = \sup\{p : E[S_T^{1+p}] < \infty\} \quad \text{et} \quad \tilde{q} = \sup\{q : E[S_T^{-q}] < \infty\}$$

4. Calculer les asymptotiques de la volatilité implicite pour T fixé et K grand/petit.

Troisième question (indépendante des deux premières) Dans cette question on s'intéresse à la couverture d'une option européenne dans le modèle (B.5).

1. Le marché, est-il complet? Peut-on construire un portefeuille de replication parfaite en utilisant une action et un autre actif liquide? Justifier votre réponse.
2. On considère la couverture approchée utilisant uniquement les actions.
   - Ecrire la dynamique du prix C_t d'une option européenne, en considérant que C_t est une fonction déterministe de S_t et v_t. On exprimera C_t comme integrale stochastique par rapport à W et à la mesure aléatoire de Poisson de Z, notée par J_Z.
   - Ecrire la dynamique d'un portefeuille autofinançant.
   - Calculer $E[\varepsilon_T^2]$ par isométrie d'Itô, où ε_T est l'erreur de couverture.
   - Calculer la stratégie de couverture quadratique optimale; montrer qu'elle coincide avec la stratégie de couverture en delta et expliquer ce résultat.

Bibliographie

[1] M. Avellaneda, R. Buff, C. Friedman, N. Grandchamp, L. Kruk, and J. Newman, *Weighted Monte Carlo: a new technique for calibrating asset-pricing models*, Int. J. Theor. Appl. Finance, 4 (2001), pp. 91–119.

[2] O. E. Barndorff-Nielsen and N. Shephard, *Econometric analysis of realized volatility and its use in estimating stochastic volatility models*, J. R. Statistic. Soc. B, 64 (2002), pp. 253–280.

[3] D. Bates, *Jumps and stochastic volatility: the exchange rate processes implicit in Deutschemark options*, Rev. Fin. Studies, 9 (1996), pp. 69–107.

[4] D. S. Bates, *Testing option pricing models*, in Statistical Methods in Finance, vol. 14 of Handbook of Statistics, North-Holland, Amsterdam, 1996, pp. 567–611.

[5] A. Ben Haj Yedder, *Calibration of stochastic volatility model with jumps*. A computer program, part of Premia software. See `www.premia.fr`, 2004.

[6] H. Berestycki, J. Busca, and I. Florent, *Asymptotics and calibration of local volatility models*, Quant. Finance, 2 (2002), pp. 61–69.

[7] T. Bollerslev, *Generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics, 31 (1986), pp. 307–327.

[8] D. Breeden and R. Litzenberger, *Prices of state-contingent claims implicit in option prices*, Journal of Business, 51 (1978), pp. 621–651.

[9] M. Broadie and O. Kaya, *Exact simulation of stochastic volatility and other affine jump diffusion processes*. Discussion paper, Columbia University, Graduate School of Business.

[10] R. Byrd, P. Lu, and J. Nocedal, *A limited memory algorithm for bound constrained optimization*, SIAM Journal on Scientific and Statistical Computing, 16 (1995), pp. 1190–1208.

[11] P. Carr and D. Madan, *Towards a theory of volatility trading*, in Volatility, R. Jarrow, ed., Risk Publications, 1998.

[12] T. F. Coleman, Y. Lee, and A. Verna, *Reconstructing the unknown local volatility function*, J. Comput. Finance, 2 (1999), pp. 77–102.

[13] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*, Chapman & Hall / CRC Press, 2004.

[14] R. Cont and P. Tankov, *Non-parametric calibration of jump-diffusion option pricing models*, J. Comput. Finance, 7 (2004), pp. 1–49.

[15] R. Cont and P. Tankov, *Retrieving Lévy processes from option prices: Regularization of an ill-posed inverse problem*, SIAM Journal on Control and Optimization, 45 (2006), pp. 1–25.

[16] R. Cont and E. Voltchkova, *Integro-differential equations for option prices in exponential Lévy models*, Finance and Stochastics, 9 (2005), pp. 299–325.

[17] J. W. Cooley and J. W. Tukey, *An algorithm for the machine calculation of complex Fourier series*, Math. Comp., 19 (1965), pp. 297–301.

[18] J. C. Cox, *The constant elasticity of variance option pricing model*, Journal of Portfolio Management, 22 (1996), pp. 15–17.

[19] J. C. Cox, J. E. Ingersoll, Jr., and S. A. Ross, *An intertemporal general equilibrium model of asset prices*, Econometrica, 53 (1985), pp. 363–384.

[20] S. Crépey, *Calibration of the local volatility in a trinomial tree using Tikhonov regularization*, Inverse Problems, 19 (2003), pp. 91–127.

[21] F. Delbaen and W. Schachermayer, *The fundamental theorem of asset pricing for unbounded stochastic processes*, Math. Ann., 312 (1998), pp. 215–250.

[22] E. Derman, *Regimes of volatility*, RISK, (1999).

[23] E. Derman and I. Kani, *Riding on a Smile*, RISK, 7 (1994), pp. 32–39.

[24] E. Derman, I. Kani, and N. Chriss, *Implied trinomial trees of the volatility smile*, The Journal of Derivatives, (Summer 1996).

[25] D. Duffie, D. Filipovic, and W. Schachermayer, *Affine processes and applications in finance*, Ann. Appl. Probab., 13 (2003), pp. 984–1053.

[26] D. Duffie, J. Pan, and K. Singleton, *Transform analysis and asset pricing for affine jump-diffusions*, Econometrica, 68 (2000), pp. 1343–1376.

[27] B. Dupire, *Pricing with a smile*, RISK, 7 (1994), pp. 18–20.

[28] N. El Karoui, *Couverture des risques dans les marchés financiers*. Lecture notes for master 'Probability and Finance', Paris VI university.

[29] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, vol. 375, Kluwer Academic Publishers Group, Dordrecht, 1996.

[30] R. F. Engle, *Autoregressive conditional heteroscedasticity with estimates of variance of united kingdom inflation*, Econometrica, 50 (1982), pp. 987–1008.

[31] P. Friz and S. Benhaim, *Regular variation and smile asymptotics*, Mathematical Finance, (to appear).

[32] M. B. Garman and S. W. Kohlhagen, *Foreign currency option values*, Journal of International Money and Finance, 2 (1983), pp. 231–237.

[33] J. Gatheral, *The Volatility Surface: a Practitioner's Guide*, Wiley Finance, 2006.

[34] H. Geman, D. Madan, and M. Yor, *Asset prices are Brownian motion: Only in business time*, in Quantitative Analysis in Financial Markets, M. Avellaneda, ed., World Scientific, River Edge, NJ, 2001, pp. 103–146.

[35] P. Glasserman, *Monte Carlo Methods in Financial Engineering*, Springer, New York, 2003.

[36] T. Goll and J. Kallsen, *Optimal portfolios for logarithmic utility*, Stochastic Process. Appl., 89 (2000), pp. 31–48.

[37] P. S. Hagan and D. E. Woodward, *Equivalent black volatilities*. Research report, 1998.

[38] J. Harrison and D. Kreps, *Martingales and arbitrage in multiperiod security markets*, J. Economic Theory, 2 (1979), pp. 381–408.

[39] T. Hayashi and P. A. Mykland, *Hedging errors: an asymptotic approach*, Mathematical Finance, 15 (2005), pp. 309–343.

[40] S. Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, Rev. Fin. Studies, 6 (1993), pp. 327–343.

[41] N. Jackson, E. Süli, and S. Howison, *Computation of deterministic volatility surfaces*, Journal of Computational Finance, 2 (1999), pp. 5–32.

[42] M. Keller-Ressel, *Moment explosions and long-term behavior of affine stochastic volatility models*. Forthcoming in Mathematical Finance, 2008.

[43] H. J. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer, 2001.

[44] R. Lagnado and S. Osher, *A technique for calibrating derivative security pricing models: numerical solution of the inverse problem*, J. Comput. Finance, 1 (1997).

[45] R. Lee, *The moment formula for implied volatility at extreme strikes*, Mathematical Finance, 14 (2004), pp. 469–480.

[46] R. Lord and C. Kahl, *Complex logarithms in heston-like models*. Available at SSRN: http://ssrn.com/abstract=1105998, 2008.

[47] D. Madan, P. Carr, and E. Chang, *The variance gamma process and option pricing*, European Finance Review, 2 (1998), pp. 79–105.

[48] D. Madan and M. Konikov, *Option pricing using variance gamma Markov chains*, Rev. Derivatives Research, 5 (2002), pp. 81–115.

[49] D. Madan and F. Milne, *Option pricing with variance gamma martingale components*, Math. Finance, 1 (1991), pp. 39–55.

[50] A. Medvedev and O. Scaillet, *A simple calibration procedure of stochastic volatility models with jumps by short term asymptotics*. Download from `www.hec.unige.ch/professeurs/SCAILLET_Olivier/pages_web/Home_Page.htm`, 2004.

[51] D. Nualart, *The Malliavin Calculus and Related Topics*, Springer, 1995.

[52] P. Protter, *Stochastic integration and differential equations*, Springer, Berlin, 1990.

[53] M. Romano and N. Touzi, *Contingent claims and market completeness in a stochastic volatility model*, Mathematical Finance, 7 (1997), pp. 399–410.

[54] M. Rubinstein, *Implied binomial trees*, Journal of Finance, 49 (1994), pp. 771–819.

[55] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge, UK, 1999.

[56] M. Stutzer, *A simple nonparametric approach to derivative security valuation*, Journal of Finance, 51 (1996), pp. 1633–1652.

[57] P. Tankov, *Lévy Processes in Finance: Inverse Problems and Dependence Modelling*, PhD thesis, Ecole Polytechnique, France, 2004.

[58] R. Zhang, *Couverture approchée des options Européennes.*, PhD thesis, Ecole Nationale des Ponts et Chaussées, 1999.